

# Pragmatic Semantic Subtyping

LILY BROWN, ANDY FRIESEN, ALAN JEFFREY, and AARON WEISS, Roblox, USA

This paper presents the view of subtyping Luau programming language. This system has been deployed as part of the Luau programming language, used by millions of users of Roblox Studio.

CCS Concepts: • **Software and its engineering** → **Semantics**.

## ACM Reference Format:

Lily Brown, Andy Friesen, Alan Jeffrey, and Aaron Weiss. 2024. Pragmatic Semantic Subtyping. *Proc. ACM Program. Lang.* 7, ICFP, Article ?? (September 2024), 16 pages. <https://doi.org/10.1145/nnnnnnnn.nnnnnnnn>

## 1 INTRODUCTION

Luau is a scripting language used by Roblox creators in the IDE tool in Roblox Studio. In 2022 there were more than 4 million creators, Fig 1 [7], which is the largest user base of Semantic Subtyping. In [1] we discuss why Luau uses Semantic Subtyping:

*Semantic subtyping interprets types as sets of values, and subtyping as set inclusion [3]. This is aligned with the minimize false positives goal of Luau non-strict mode, since semantic subtyping only reports a failure of subtyping when there is a value which inhabits the candidate subtype, but not the candidate supertype. For example, the program:*

```
local x : CFrame
    = CFrame.new()
local y : Vector3 | CFrame
    = (math.random() < 0.5 ? CFrame.new() : Vector3.new())
local z : Vector3 | CFrame
    = x * y
```

*cannot produce a run-time error, since multiplication of CFrames is overloaded:*

```
((CFrame, CFrame) -> CFrame) & ((CFrame, Vector3) -> Vector3)
```

*In order to typecheck this program, we check that that type is a subtype of:*

```
(CFrame, Vector3 | CFrame) -> (Vector3 | CFrame)
```

*In the previous, syntax-driven, implementation of subtyping, this subtype check would fail, resulting in a false positive. We have now released an implementation of semantic subtyping, which does not suffer from this defect. See our technical blog for more details [5].*

In Luau, we use a variant of semantic subtyping [3, 4, 6]. The important properties of semantic subtyping are:

- there is a set  $\mathcal{D}$  of semantic values,
- each type  $T$  has a semantics  $\llbracket T \rrbracket \subseteq \mathcal{D}$ ,
- unknown and never types are interpreted as  $\mathcal{D}$  and  $\emptyset$ ,
- union and intersection types are interpreted as set union and intersection, and
- subtyping  $T <: U$  is interpreted as  $\llbracket T \rrbracket \subseteq \llbracket U \rrbracket$ .

---

Authors' address: Lily Brown; Andy Friesen; Alan Jeffrey; Aaron Weiss, Roblox, San Mateo, CA, USA.



This work is licensed under a Creative Commons Attribution 4.0 International License.

© 2024 Roblox.

2475-1421/2024/9-ART?? \$15.00

<https://doi.org/10.1145/nnnnnnnn.nnnnnnnn>



Fig. 1. Creators numbers in 2022

The off-the-shelf presentation of semantic subtyping is *set theoretic* [3, §2.5]:

$$\llbracket T_1 \rrbracket \subseteq \llbracket T_2 \rrbracket \text{ if and only if } \mathcal{E}[\llbracket T_1 \rrbracket] \subseteq \mathcal{E}[\llbracket T_2 \rrbracket]$$

where the most important case for  $\mathcal{E}[\llbracket T \rrbracket]$  is function types:

$$\mathcal{E}[\llbracket S \rightarrow T \rrbracket] = \mathcal{P}(\mathcal{D}^2 \setminus (\llbracket S \rrbracket \times (\mathcal{D} \setminus \llbracket T \rrbracket)))$$

The set theoretical requirement has some consequences:

*All functions types (never  $\rightarrow T$ ) are identified. Consider*

$$\begin{aligned} \mathcal{E}[\llbracket \text{never} \rightarrow T_1 \rrbracket] &= \mathcal{P}(\mathcal{D}^2 \setminus (\llbracket \text{never} \rrbracket \times (\mathcal{D} \setminus \llbracket T_1 \rrbracket))) \\ &= \mathcal{P}(\mathcal{D}^2 \setminus (\emptyset \times (\mathcal{D} \setminus \llbracket T_1 \rrbracket))) \\ &= \mathcal{P}(\mathcal{D}^2) \\ &= \mathcal{P}(\mathcal{D}^2 \setminus (\emptyset \times (\mathcal{D} \setminus \llbracket T_2 \rrbracket))) \\ &= \mathcal{P}(\mathcal{D}^2 \setminus (\llbracket \text{never} \rrbracket \times (\mathcal{D} \setminus \llbracket T_2 \rrbracket))) \\ &= \mathcal{E}[\llbracket \text{never} \rightarrow T_2 \rrbracket] \end{aligned}$$

in particular, this means we cannot define a semantics-preserving function  $\text{apply}(T, U)$  such that:

$$\text{apply}(S \rightarrow T, U) = T \text{ when } U <: S$$

because there is a nasty case where  $S$  is uninhabited. In this presentation, the  $\text{apply}$  function used in the rule for function application:

$$\frac{D_1 : (\Gamma \vdash M : T) \quad D_2 : (\Gamma \vdash N : U)}{\text{app}(D_1, D_2) : (\Gamma \vdash M(N) : \text{apply}(T, U))}$$

so we have to accept that in a set-theoretic model, the type rule for function application has corner cases for uninhabited types.

*Union does not distributed through function types.* Semantic subtyping gives a natural model of overloaded functions as intersections of arrows, for example the Roblox API for matrices include an overloaded function which supports multiplication of both 2D (CFrame) and 1D (Vector3) matrices:

`CFrame.__mul : ((CFrame, CFrame) -> CFrame) & ((CFrame, Vector3) -> Vector3)`

Overloaded functions are a key part of the Roblox API, and we might expect that all function types can be presented as overloaded functions (that is intersections of arrows). We can do that it we can

$$\begin{aligned}
v &::= s \mid (a \mapsto r) \\
a &::= () \mid (v) \\
r &::= \text{diverge} \mid \text{check} \mid (v)
\end{aligned}$$

Fig. 2. Semantic values

distribute union through arrow:

$$\llbracket (S_1 \rightarrow T_1) \cup (S_2 \rightarrow T_2) \rrbracket = \llbracket (S_1 \cap S_2) \rightarrow (T_1 \cup T_2) \rrbracket$$

For example:

$$\llbracket (\text{number?} \rightarrow \text{number}) \cup (\text{string?} \rightarrow \text{string}) \rrbracket = \llbracket \text{nil} \rightarrow (\text{number} \cup \text{string}) \rrbracket$$

Unfortunately, set-theoretic models do not allow union to distributed through intersection, for example:

$$\begin{aligned}
\{(1, \text{nil}), (\text{"hi"}, \text{nil})\} &\in \mathcal{E} \llbracket \text{nil} \rightarrow (\text{number} \cup \text{string}) \rrbracket \\
\{(1, \text{nil}), (\text{"hi"}, \text{nil})\} &\notin \mathcal{E} \llbracket \text{number?} \rightarrow \text{number} \rrbracket \\
\{(1, \text{nil}), (\text{"hi"}, \text{nil})\} &\notin \mathcal{E} \llbracket \text{string?} \rightarrow \text{string} \rrbracket
\end{aligned}$$

This is why type normalization for function types in set-theoretic models uses a conjunctive normal form of unions of intersections of functions e.g. [6, §4.1.2].

*Set-theoretic mode support negated types.* In addition, Luau does not support negation of all types, but only negation of *test types* [?], which simplifies the model, by not requiring arbitrary type negation. In particular, since the model does not support negation of function types, the normal form for function types is just overload functions, not combinations of positive and negative function types.

*Conclusions of this paper.* In summary there is a trade-off in semantic subtyping:

- *set-theoretic* models, which are closer to the set-theoretic model of functions, and
- *pragmatic* models, which drop the set-theoretic requirement, and in return a) do not have corner cases on the type of function application when the argument has uninhabited type, and b) have overloaded functions (that is intersections of arrows) as the normal for function types.

Luau chooses to adopt a pragmatic semantic subtyping model.

This paper shows how core Luau pragmatic model is defined, and how it formally (in Agda) proves pragmatic models. There is the in the the full Luau which is much bigger than the formally core language.

## 2 FORMAL TREATMENT OF CORE LUAU

This is a formal of a small core language. It has scalar types (`nil`, `number`, `boolean` and `string`), union and intersection types (for example the optional  $T?$  is a common shorthand  $T \cup \text{nil}$ ), and single-arity functions (like `number?  $\rightarrow$  number`). In this section why the core language can be formally, and in particularly type normalization proved a algorithm for checking subtyping.

### 2.1 Semantic Values for Core Luau

In this presentation, we present the minimal core of Luau, which supports scalars and functions. This presentation ignores tables, mutable features, and object objects. We will ignore the details of scalar types, and assume that there are scalar types, ranged over by  $s$ , such as `nil`, `boolean`, `number` and `string`.

The types we are considering are:

$$S, T ::= s \mid \text{unknown} \mid \text{never} \mid S \rightarrow T \mid S \cap T \mid S \cup T$$

which are:

- the *scalar types*  $s$
- the *top type*  $\text{unknown}$ ,
- the *bottom type*  $\text{never}$ ,
- a *function type*  $S \rightarrow T$ ,
- an *intersection type*  $S \cap T$ , and
- a *union type*  $S \cup T$ .

To give a semantic subtyping, we first declare the domain  $\mathcal{D}$  of *semantic values*, given by the grammar  $v$  of Figure 2. Semantic values are:

- *scalar values*  $s$ , and
- *function values*  $a \mapsto r$ , modeling a function that can map an argument  $a$  to a result  $r$ .

For example:

- $\text{true}$  and  $\text{false}$  are values in  $\text{boolean}$ ,
- $\text{true}$  and  $\text{false}$  and  $\text{nil}$  are values in the optional type  $\text{boolean} \cup \text{nil}$ , and
- $(\text{true}) \mapsto (\text{false})$  is a value in the function type  $\text{boolean} \rightarrow \text{boolean}$ .

Scalar and error-suppressing values are relatively straightforward, but functions are trickier. The case where a type-correct argument is supplied and a type-correct result is returned is clean, for example:

$$((\text{true}) \mapsto (\text{false})) \in \llbracket \text{boolean} \rightarrow \text{boolean} \rrbracket$$

But there is also the case where a type-incorrect argument is supplied, in which case there is no guarantee what is returned, for example:

$$((5) \mapsto (37)) \in \llbracket \text{boolean} \rightarrow \text{boolean} \rrbracket$$

The type-correctness guarantee for results applies when a type-correct argument is provided:

$$((\text{true}) \mapsto (37)) \notin \llbracket \text{boolean} \rightarrow \text{boolean} \rrbracket$$

Those examples consider cases where one value is supplied as an argument, and one is returned, but Luau allows other cases. Luau, as is common in most functional languages, allows functions to diverge (modeled in this semantics as  $a \mapsto \text{diverge}$ ), for example:

$$((\text{true}) \mapsto \text{diverge}) \in \llbracket \text{boolean} \rightarrow \text{boolean} \rrbracket$$

and:

$$((5) \mapsto \text{diverge}) \in \llbracket \text{boolean} \rightarrow \text{boolean} \rrbracket$$

Luau allows functions to check arguments, (modeled in this semantics as  $a \mapsto \text{check}$  when a checked fails), for example:

$$((5) \mapsto \text{check}) \in \llbracket \text{boolean} \rightarrow \text{boolean} \rrbracket$$

but:

$$((\text{true}) \mapsto \text{check}) \notin \llbracket \text{boolean} \rightarrow \text{boolean} \rrbracket$$

Luau allows functions to be called without any arguments (modeled in this semantics as  $() \mapsto r$ ) for example:

$$(() \mapsto (\text{false})) \in \llbracket \text{boolean} \rightarrow \text{boolean} \rrbracket$$

and:

$$(() \mapsto \text{diverge}) \in \llbracket \text{boolean} \rightarrow \text{boolean} \rrbracket$$

$$\begin{aligned}
\llbracket s \rrbracket &= \{s\} \\
\llbracket \text{unknown} \rrbracket &= \mathcal{D} \\
\llbracket \text{never} \rrbracket &= \emptyset \\
\llbracket S \rightarrow T \rrbracket &= \{a \mapsto (w) \mid w \in \llbracket T \rrbracket\} \cup \\
&\quad \{(v) \mapsto r \mid v \in \llbracket S \rrbracket^C\} \cup \\
&\quad \{a \mapsto \text{diverge}\} \cup \\
&\quad \{() \mapsto \text{check}\} \\
\llbracket S \cap T \rrbracket &= \llbracket S \rrbracket \cap \llbracket T \rrbracket \\
\llbracket S \cup T \rrbracket &= \llbracket S \rrbracket \cup \llbracket T \rrbracket
\end{aligned}$$

Fig. 3. Semantics of types as sets of values

and:

$$(() \mapsto \text{check}) \in \llbracket \text{boolean} \rightarrow \text{boolean} \rrbracket$$

The restriction on zero-argument function calls is that they are allowed to return a check (since they have been passed the wrong number of arguments) but they are not just allowed to return arbitrary nonsense:

$$(() \mapsto (5)) \notin \llbracket \text{boolean} \rightarrow \text{boolean} \rrbracket$$

At this point we have introduced the semantic values used by the Luau type system, and can turn the semantics of types, from which semantic subtyping follows.

## 2.2 Semantics of Core Luau Types

The semantics of Luau types are given in Fig 3. This semantics is presented mechanically in Agda in [2], and we will give the most important results here.

For example, two of the important rules are for functions, in the case where functions are called with argument values, and return result values. The rules are:

- **Type-incorrect argument:** if  $v \notin \llbracket S \rrbracket$  then  $((v) \mapsto r) \in \llbracket S \rightarrow T \rrbracket$
- **Type-correct result:** if  $w \in \llbracket T \rrbracket$  then  $(a \mapsto (w)) \in \llbracket S \rightarrow T \rrbracket$

This is the same as the semantics of Coppo types [?] as used in the fully abstract semantics of Lazy Lambda Calculus [?] using Domain Theory In Logical Form [?].

$$((v) \mapsto (w)) \in \llbracket S \rightarrow T \rrbracket \text{ if and only if } (v \in \llbracket S \rrbracket) \Rightarrow (w \in \llbracket T \rrbracket)$$

In order to give a constructive presentation of the semantics, rather than usual negative presentation of  $v \notin \llbracket S \rrbracket$ , we give a positive presentation  $v \in \llbracket S \rrbracket^C$ , as given in Fig 4.

It is routine to check that  $\llbracket S \cap T \rrbracket^C$  is a constructive presentation of  $\mathcal{D} \setminus \llbracket S \cap T \rrbracket$ .

LEMMA 2.1.  $(v \in \llbracket T \rrbracket^C)$  if and only if  $(v \notin \llbracket T \rrbracket)$ .

PROOF. An proof by injunction on  $T$  showings that  $\llbracket T \rrbracket$  is the negative of  $\llbracket T \rrbracket^C$ . □

Moreover there is a decision procedure for  $v \in \llbracket T \rrbracket$  or  $v \in \llbracket T \rrbracket^C$ .

LEMMA 2.2.  $(v \in \llbracket T \rrbracket) \vee (v \in \llbracket T \rrbracket^C)$ .

PROOF. An proof by injunction on  $T$ , that for any  $v$ , either  $v \in \llbracket T \rrbracket$  or  $v \in \llbracket T \rrbracket^C$ . □

$$\begin{aligned}
 \llbracket s \rrbracket^C &= \{t \mid s \neq t\} \\
 \llbracket \text{unknown} \rrbracket^C &= \emptyset \\
 \llbracket \text{never} \rrbracket^C &= \mathcal{D} \\
 \llbracket S \rightarrow T \rrbracket^C &= \{s\} \cup \\
 &\quad \{( ) \mapsto (w) \mid w \in \llbracket T \rrbracket^C\} \cup \\
 &\quad \{(v) \mapsto (w) \mid v \in \llbracket T \rrbracket \mid w \in \llbracket T \rrbracket^C\} \cup \\
 &\quad \{(v) \mapsto \text{check} \mid v \in \llbracket T \rrbracket\} \\
 \llbracket S \cap T \rrbracket^C &= \llbracket S \rrbracket^C \cup \llbracket T \rrbracket^C \\
 \llbracket S \cup T \rrbracket^C &= \llbracket S \rrbracket^C \cap \llbracket T \rrbracket^C
 \end{aligned}$$

Fig. 4. Complemented semantics of types as sets of values

### 2.3 Properties of Semantic Subtyping

From the semantics of types as sets of semantic values, semantic subtyping  $S <: T$  is a proof than if  $v \in \llbracket S \rrbracket$ , then  $v \in \llbracket T \rrbracket$ . Constructively this a dependent function:

$$(S <: T) \text{ if and only if } \forall v . (v \in \llbracket S \rrbracket) \rightarrow (v \in \llbracket T \rrbracket)$$

for example  $\text{number} <: \text{number?}$  since:

$$\forall v . (v \in \llbracket \text{number} \rrbracket) \rightarrow (v \in \llbracket \text{number?} \rrbracket)$$

Subtyping can be viewed as a dependent function from  $\llbracket T \rrbracket^C$  to  $\llbracket S \rrbracket^C$

LEMMA 2.3.  $(S <: T)$  if and only if  $\forall v . (v \in \llbracket T \rrbracket^C) \rightarrow (v \in \llbracket S \rrbracket^C)$

PROOF. For “if”, for any  $v$ , if  $v \in \llbracket S \rrbracket$ , then by Lemma 2.2 either  $v \in \llbracket T \rrbracket$  or  $v \in \llbracket T \rrbracket^C$ . In the first case,  $v \in \llbracket T \rrbracket$  as needed. In the second case  $v \in \llbracket T \rrbracket^C$  and so by “if” hypothesis  $v \in \llbracket S \rrbracket^C$ , but by Lemma 2.1 we have a contradiction from  $v \in \llbracket S \rrbracket$  and  $v \in \llbracket S \rrbracket^C$ . So we have established that  $S <: T$ .

For “only if”, for any  $v$ , if  $v \in \llbracket T \rrbracket^C$ , then by Lemma 2.2 either  $v \in \llbracket S \rrbracket$  or  $v \in \llbracket S \rrbracket^C$ . In the first case  $v \in \llbracket S \rrbracket$  and so by “only if” hypothesis,  $v \in \llbracket T \rrbracket$ , but by Lemma 2.1 we have a contradiction from  $v \in \llbracket T \rrbracket$  and  $v \in \llbracket T \rrbracket^C$ . In the second case,  $v \in \llbracket S \rrbracket^C$  as needed. So we have established that  $\forall v . (v \in \llbracket T \rrbracket^C) \rightarrow (v \in \llbracket S \rrbracket^C)$ .  $\square$

More interestingly is the constructive presentation of *anti-subtyping*  $S \nless T$ . Normally this is presented negatively, but it can be read constructively since  $S \nless T$  is witnessed by a value  $v$  where  $v \in \llbracket S \rrbracket$  but  $v \in \llbracket T \rrbracket^C$ .

$$(S \nless T) \text{ if and only if } \exists v . (v \in \llbracket S \rrbracket) \wedge (v \in \llbracket T \rrbracket^C)$$

for example  $\text{number?} \nless \text{number}$  since we can pick our witness  $v$  to be  $\text{nil}$ :

$$\text{nil} \in \llbracket \text{number?} \rrbracket \quad \text{nil} \in \llbracket \text{number} \rrbracket^C$$

Now, by Lemma 2.1, it is direct that  $S \nless T$  is a contradiction of  $S <: T$ :

LEMMA 2.4.  $(S \nless T) \rightarrow \neg(S <: T)$

PROOF. The is a witness for  $S \nless T$ ,  $v \in \llbracket S \rrbracket$  and  $v \in \llbracket T \rrbracket^C$ , and so  $S <: T$  and  $v \in \llbracket S \rrbracket$  gives  $v \in \llbracket T \rrbracket$ . So Lemma 2.1 gives  $v \in \llbracket T \rrbracket^C$  and  $v \in \llbracket T \rrbracket$  are a contradiction is required.  $\square$

Unfortunately, this does not give a decision procedure for subtyping, for the usual reason that it is tricky to build an algorithm for checking semantic subtyping, which requires type normalization [?]. We will return to this in §2.6.

It is direct to show that  $<$  is transitive.

LEMMA 2.5.  $(S < T) \wedge (T < U) \rightarrow (S < U)$

PROOF. There must be  $f$  maps an argument with type  $v \in \llbracket S \rrbracket$  to a result with type  $v \in \llbracket T \rrbracket$ , and there must be  $g$  maps an argument with type  $v \in \llbracket T \rrbracket$  to a result with type  $v \in \llbracket U \rrbracket$ . So  $f;g$  maps an argument with type  $v \in \llbracket S \rrbracket$  to a result with type  $v \in \llbracket U \rrbracket$ .  $\square$

More interestingly there is a dual property for  $\nless$ . Classically this is the same as transitivity, just stated in terms of  $\nless$  rather than  $<$ . But constructively this is a choice function, that states that if  $S \nless U$  then for any  $T$  we have a witness for either  $S \nless T$  or  $T \nless U$ . For example  $\text{number?} \nless \text{number}$  (witnessed by  $\text{nil}$ ), so for a mid-point string we have  $\text{nil} \in \llbracket \text{string} \rrbracket^C$  which means we chose the constructive anti-subtype  $\text{number?} \nless \text{string}$  witnessed by  $\text{nil} \in \llbracket \text{number?} \rrbracket$  and  $\text{nil} \in \llbracket \text{string} \rrbracket^C$ .

LEMMA 2.6.  $(S \nless U) \rightarrow (S \nless T) \vee (T \nless U)$

PROOF.  $(S \nless U)$  must have a witness  $v$  where  $v \in \llbracket S \rrbracket$  and  $v \in \llbracket U \rrbracket^C$ . Now by Lemma 2.2 (the decision procedure for type semantics) we either have  $v \in \llbracket T \rrbracket$  or  $v \in \llbracket T \rrbracket^C$ . In the first case,  $v \in \llbracket T \rrbracket$  and  $v \in \llbracket U \rrbracket^C$ , which witnesses  $T \nless U$ . In the second case,  $v \in \llbracket S \rrbracket$  and  $v \in \llbracket T \rrbracket^C$ , which witnesses  $S \nless T$ . In either case, we have a decision procedure for  $(S \nless T) \vee (T \nless U)$ .  $\square$

## 2.4 Co- and Contra-variance of Subtyping for Functions

We now turn to co- and contra-variant subtyping of functions. These come in two flavors: when function types are introduced, and when function types are eliminated. When a function type is introduced, we check that the arguments respect contravariant subtyping, and that the results respect covariant subtyping.

LEMMA 2.7. *If  $S' < S$  and  $T < T'$  then  $(S \rightarrow T) < (S' \rightarrow T')$*

PROOF. If  $u \in \llbracket S \rightarrow T \rrbracket$  then from Fig 3, either

- $u = (a \mapsto (w))$  and  $w \in \llbracket T \rrbracket$ , so  $T < T'$  implies  $w \in \llbracket T' \rrbracket$ , and so  $(a \mapsto (w)) \in \llbracket S' \rightarrow T' \rrbracket$ ,
- $u = ((v) \mapsto r)$  and  $v \in \llbracket S \rrbracket^C$ , so  $S' < S$  and Lemma 2.3 implies  $v \in \llbracket S' \rrbracket^C$ , and so  $((v) \mapsto r) \in \llbracket S' \rightarrow T' \rrbracket$ ,
- $u = (a \mapsto \text{diverge})$ , so  $(a \mapsto \text{diverge}) \in \llbracket S' \rightarrow T' \rrbracket$ , or
- $u = () \mapsto \text{check}$ , so  $(() \mapsto \text{check}) \in \llbracket S' \rightarrow T' \rrbracket$ .

In any case,  $u \in \llbracket S' \rightarrow T' \rrbracket$ .  $\square$

When a function type is eliminated, we check that the arguments reflect contravariant subtyping, and that the results reflect covariant subtyping.

LEMMA 2.8. *If  $(S \rightarrow T) < (S' \rightarrow T')$  then  $S' < S$  and  $T < T'$*

PROOF. If  $v \in \llbracket S \rrbracket^C$  then from Fig 3  $((v) \mapsto \text{check}) \in \llbracket S \rightarrow T \rrbracket$ , so since  $(S \rightarrow T) < (S' \rightarrow T')$ , we have  $((v) \mapsto \text{check}) \in \llbracket S' \rightarrow T' \rrbracket$  and so  $v \in \llbracket S' \rrbracket^C$ . Thus, using Lemma 2.3, we have established  $S' < S$ .

If  $w \in \llbracket T \rrbracket$  then from Fig 3  $(() \mapsto (w)) \in \llbracket S \rightarrow T \rrbracket$ , so since  $(S \rightarrow T) < (S' \rightarrow T')$ , we have  $(() \mapsto (w)) \in \llbracket S' \rightarrow T' \rrbracket$  and so  $w \in \llbracket T' \rrbracket$ . Thus we have established  $T < T'$ .  $\square$

Note that these Lemmas rely on pragmatic semantic subtyping. Lemma 2.7 is true for set-theoretic semantic subtyping, but Lemma 2.8 is only true for set-theoretic models when  $S$  and  $T$  are inhabited types. In pragmatic semantic subtyping, we do not have special corner cases, in particular Lemma 2.8 is true for all types, and does not require special cases about inhabitation.

## 2.5 Distribution of Intersection and Union Over Functions

Finally, we turn to cases where intersection and union distribute through functions. Since Luau uses intersections of functions as types for overloaded functions, it is unsurprising that intersection does not in general distribute through functions. For example:

$$(\text{boolean} \rightarrow \text{boolean}) \cap (\text{number} \rightarrow \text{number})$$

does not distribute to:

$$(\text{boolean} \cup \text{number}) \rightarrow (\text{boolean} \cap \text{number})$$

For example:

$$\begin{aligned} (\text{false} \mapsto \text{true}) &\in \llbracket (\text{boolean} \rightarrow \text{boolean}) \cap (\text{number} \rightarrow \text{number}) \rrbracket \\ (\text{false} \mapsto \text{true}) &\in \llbracket (\text{boolean} \cup \text{number}) \rightarrow (\text{boolean} \cap \text{number}) \rrbracket^C \end{aligned}$$

is a witness for:

$$\begin{aligned} &((\text{boolean} \rightarrow \text{boolean}) \cap (\text{number} \rightarrow \text{number})) \\ &\not\leq ((\text{boolean} \cup \text{number}) \rightarrow (\text{boolean} \cap \text{number})) \end{aligned}$$

Now in general we can not distribute intersection through functions, on both the left and right, but we can distribute on just left, and just right. This is similar to the situation with premonoids categories  $[?]^\perp$  which are functorial on both sides, but are not binary functorials.

LEMMA 2.9.

- (1)  $\llbracket (S_1 \rightarrow T) \cap (S_2 \rightarrow T) \rrbracket = \llbracket (S_1 \cup S_2) \rightarrow T \rrbracket$
- (2)  $\llbracket (S \rightarrow T_1) \cap (S \rightarrow T_2) \rrbracket = \llbracket S \rightarrow (T_1 \cap T_2) \rrbracket$

PROOF.

(1 $\Rightarrow$ ) If  $u \in \llbracket S_1 \rightarrow T \rrbracket \cap \llbracket S_2 \rightarrow T \rrbracket$ , then from Fig 3, either

- $u = (a \mapsto (w))$  and  $w \in \llbracket T \rrbracket$ , and so  $(a \mapsto (w)) \in \llbracket (S_1 \cup S_2) \rightarrow T \rrbracket$ ,
- $u = ((v) \mapsto r)$ ,  $v \in \llbracket S_1 \rrbracket^C$  and  $v \in \llbracket S_2 \rrbracket^C$ , so  $v \in \llbracket S_1 \cup S_2 \rrbracket^C$ , and so  $((v) \mapsto r) \in \llbracket (S_1 \cup S_2) \rightarrow T \rrbracket$ ,
- $u = (a \mapsto \text{diverge})$ , so  $(a \mapsto \text{diverge}) \in \llbracket (S_1 \cup S_2) \rightarrow T \rrbracket$ , or
- $u = () \mapsto \text{check}$ , so  $(() \mapsto \text{check}) \in \llbracket (S_1 \cup S_2) \rightarrow T \rrbracket$ .

In any case,  $u \in \llbracket (S_1 \cup S_2) \rightarrow T \rrbracket$ .

(1 $\Leftarrow$ ) If  $u \in \llbracket (S_1 \cup S_2) \rightarrow T \rrbracket$ , then from Fig 3, either

- $u = (a \mapsto (w))$  and  $w \in \llbracket T \rrbracket$ , and so  $(a \mapsto (w)) \in \llbracket S_1 \rightarrow T \rrbracket \cap \llbracket S_2 \rightarrow T \rrbracket$ ,
- $u = ((v) \mapsto r)$ ,  $v \in \llbracket S_1 \rrbracket^C \cap \llbracket S_2 \rrbracket^C$  and so  $((v) \mapsto r) \in \llbracket S_1 \rightarrow T \rrbracket \cap \llbracket S_2 \rightarrow T \rrbracket$ ,
- $u = (a \mapsto \text{diverge})$ , so  $(a \mapsto \text{diverge}) \in \llbracket S_1 \rightarrow T \rrbracket \cap \llbracket S_2 \rightarrow T \rrbracket$ , or
- $u = () \mapsto \text{check}$ , so  $(() \mapsto \text{check}) \in \llbracket S_1 \rightarrow T \rrbracket \cap \llbracket S_2 \rightarrow T \rrbracket$ .

In any case,  $u \in \llbracket S_1 \rightarrow T \rrbracket \cap \llbracket S_2 \rightarrow T \rrbracket$ .

(2 $\Rightarrow$ ) If  $u \in \llbracket S \rightarrow T_1 \rrbracket \cap \llbracket S \rightarrow T_2 \rrbracket$ , then from Fig 3, either

- $u = (a \mapsto (w))$ ,  $w \in \llbracket T_1 \rrbracket$  and  $w \in \llbracket T_2 \rrbracket$ , and so  $(a \mapsto (w)) \in \llbracket S \rightarrow (T_1 \cap T_2) \rrbracket$ ,
- $u = ((v) \mapsto r)$  and  $v \in \llbracket S \rrbracket^C$  so  $v \in \llbracket S \rrbracket^C$ , and so  $((v) \mapsto r) \in \llbracket S \rightarrow (T_1 \cap T_2) \rrbracket$ ,
- $u = (a \mapsto \text{diverge})$ , so  $(a \mapsto \text{diverge}) \in \llbracket S \rightarrow (T_1 \cap T_2) \rrbracket$ , or
- $u = () \mapsto \text{check}$ , so  $(() \mapsto \text{check}) \in \llbracket S \rightarrow (T_1 \cap T_2) \rrbracket$ .

In any case,  $u \in \llbracket S \rightarrow (T_1 \cap T_2) \rrbracket$ .



$$\begin{aligned}
F &::= \text{never} \mid (S_1 \rightarrow T_1) \cap \dots \cap (S_m \rightarrow T_m) \\
N &::= F \cup s_1 \cup \dots \cup s_n
\end{aligned}$$

Fig. 5. Normalized function types (ranged over by  $F$ ) and types (ranged over by  $N$ )

( $2 \Rightarrow$ ) If  $u \in \llbracket (S_1 \cup S_2) \rightarrow T \rrbracket$ , then from Fig 3, either

- $u = (a \mapsto (w))$ ,  $w \in \llbracket T_1 \rrbracket$  and  $w \in \llbracket T_2 \rrbracket$ , and so  $(a \mapsto (w)) \in \llbracket S \rightarrow T_1 \rrbracket \cap \llbracket S \rightarrow T_2 \rrbracket$ ,
- $u = ((v) \mapsto r)$  and  $v \in \llbracket S \rrbracket^C$  so  $v \in \llbracket S \rrbracket^C$ , and so  $((v) \mapsto r) \in \llbracket S \rightarrow T_1 \rrbracket \cap \llbracket S \rightarrow T_2 \rrbracket$ ,
- $u = (a \mapsto \text{diverge})$ , so  $(a \mapsto \text{diverge}) \in \llbracket S \rightarrow T_1 \rrbracket \cap \llbracket S \rightarrow T_2 \rrbracket$ , or
- $u = () \mapsto \text{check}$ , so  $(() \mapsto \text{check}) \in \llbracket S \rightarrow T_1 \rrbracket \cap \llbracket S \rightarrow T_2 \rrbracket$ .

In any case,  $u \in \llbracket S \rightarrow T_1 \rrbracket \cap \llbracket S \rightarrow T_2 \rrbracket$ .

This provides cases. □

In contract, union does distribute over functions.

LEMMA 2.10.  $\llbracket (S_1 \rightarrow T_1) \cup (S_2 \rightarrow T_2) \rrbracket = \llbracket (S_1 \cap S_2) \rightarrow (T_1 \cup T_2) \rrbracket$

PROOF.

( $\Rightarrow$ ) If  $u \in \llbracket S_1 \rightarrow T_1 \rrbracket \cup \llbracket S_2 \rightarrow T_2 \rrbracket$ , then from Fig 3, either

- $u = (a \mapsto (w))$ ,  $w \in \llbracket T_1 \rrbracket$ , and so  $(a \mapsto (w)) \in \llbracket (S_1 \cap S_2) \rightarrow (T_1 \cup T_2) \rrbracket$ ,
- $u = (a \mapsto (w))$ ,  $w \in \llbracket T_2 \rrbracket$ , and so  $(a \mapsto (w)) \in \llbracket (S_1 \cap S_2) \rightarrow (T_1 \cup T_2) \rrbracket$ ,
- $u = ((v) \mapsto r)$ ,  $v \in \llbracket S_1 \rrbracket^C$  and so  $((v) \mapsto r) \in \llbracket (S_1 \cap S_2) \rightarrow (T_1 \cup T_2) \rrbracket$ ,
- $u = ((v) \mapsto r)$ ,  $v \in \llbracket S_2 \rrbracket^C$  and so  $((v) \mapsto r) \in \llbracket (S_1 \cap S_2) \rightarrow (T_1 \cup T_2) \rrbracket$ ,
- $u = (a \mapsto \text{diverge})$ , so  $(a \mapsto \text{diverge}) \in \llbracket (S_1 \cap S_2) \rightarrow (T_1 \cup T_2) \rrbracket$ , or
- $u = () \mapsto \text{check}$ , so  $(() \mapsto \text{check}) \in \llbracket (S_1 \cap S_2) \rightarrow (T_1 \cup T_2) \rrbracket$ .

In any case,  $u \in \llbracket (S_1 \cap S_2) \rightarrow (T_1 \cup T_2) \rrbracket$ .

( $\Leftarrow$ ) If  $u \in \llbracket (S_1 \cap S_2) \rightarrow (T_1 \cup T_2) \rrbracket$ , then from Fig 3, either

- $u = (a \mapsto (w))$ ,  $w \in \llbracket T_1 \rrbracket$ , and so  $(a \mapsto (w)) \in \llbracket S_1 \rightarrow T_1 \rrbracket \cup \llbracket S_2 \rightarrow T_2 \rrbracket$ ,
- $u = (a \mapsto (w))$ ,  $w \in \llbracket T_2 \rrbracket$ , and so  $(a \mapsto (w)) \in \llbracket S_1 \rightarrow T_1 \rrbracket \cup \llbracket S_2 \rightarrow T_2 \rrbracket$ ,
- $u = ((v) \mapsto r)$ ,  $v \in \llbracket S_1 \rrbracket^C$  and so  $((v) \mapsto r) \in \llbracket S_1 \rightarrow T_1 \rrbracket \cup \llbracket S_2 \rightarrow T_2 \rrbracket$ ,
- $u = ((v) \mapsto r)$ ,  $v \in \llbracket S_2 \rrbracket^C$  and so  $((v) \mapsto r) \in \llbracket S_1 \rightarrow T_1 \rrbracket \cup \llbracket S_2 \rightarrow T_2 \rrbracket$ ,
- $u = (a \mapsto \text{diverge})$ , so  $(a \mapsto \text{diverge}) \in \llbracket S_1 \rightarrow T_1 \rrbracket \cup \llbracket S_2 \rightarrow T_2 \rrbracket$ , or
- $u = () \mapsto \text{check}$ , so  $(() \mapsto \text{check}) \in \llbracket S_1 \rightarrow T_1 \rrbracket \cup \llbracket S_2 \rightarrow T_2 \rrbracket$ .

In any case,  $u \in \llbracket S_1 \rightarrow T_1 \rrbracket \cup \llbracket S_2 \rightarrow T_2 \rrbracket$ .

This provides cases. □

Now, we have that the Luau pragmatic semantic subtyping unions distinction, the type normal for functions is:

$$(S_1 \rightarrow T_1) \cap \dots \cap (S_N \rightarrow T_N)$$

This the most important feature of type normalization.

## 2.6 Type Normalization

For Core Luau, type normalization is simplifier compared than sub-theoretic semantic subtyping, as shown in Fig 6. A normalized function type is a overloaded function, for example:

$$(\text{string?} \rightarrow \text{number?}) \cap (\text{number} \rightarrow \text{number})$$

The difficult part of normalized types is function types, but we also consider normalized types with scalars for example optional number

$$\text{number} \cup \text{nil}$$

or a function type that may be a function type or a scalar type for example the overloaded function or optional number:

$$((\text{string?} \rightarrow \text{number?}) \cap (\text{number} \rightarrow \text{number})) \cup \text{number} \cup \text{nil}$$

LEMMA 2.11. *For any normalized function type  $F$  and  $F'$ , there is one equivalent to  $\llbracket F \rrbracket \cup \llbracket F' \rrbracket$ .*

PROOF. If  $F$  is never then  $\llbracket F' \rrbracket$  equivalent to  $\llbracket F \rrbracket \cup \llbracket F' \rrbracket$ . If  $F'$  is never then  $\llbracket F \rrbracket$  equivalent to  $\llbracket F \rrbracket \cup \llbracket F' \rrbracket$ . Otherwise:

$$F = (S_1 \rightarrow T_1) \cap \dots \cap (S_m \rightarrow T_m) \quad F' = (S'_1 \rightarrow T'_1) \cap \dots \cap (S'_{m'} \rightarrow T'_{m'})$$

We first distribution union through intersect:

$$((S_1 \rightarrow T_1) \cap \dots \cap (S_m \rightarrow T_m)) \cup ((S'_1 \rightarrow T'_1) \cap \dots \cap (S'_{m'} \rightarrow T'_{m'}))$$

to get an equivalent non-normal type:

$$((S_1 \rightarrow T_1) \cup (S'_1 \rightarrow T'_1)) \cap \dots \cap ((S_1 \rightarrow T_1) \cup (S'_{m'} \rightarrow T'_{m'}))$$

$$\cap$$

$$\vdots$$

$$\cap$$

$$((S_m \rightarrow T_m) \cup (S'_1 \rightarrow T'_1)) \cap \dots \cap ((S_m \rightarrow T_m) \cup (S'_{m'} \rightarrow T'_{m'}))$$

Now we use Lemma 2.10 to distribute union through function:

$$((S_1 \cap S'_1) \rightarrow (T_1 \cup T'_1)) \cap \dots \cap ((S_1 \cap S'_{m'}) \rightarrow (T_1 \cup T'_{m'}))$$

$$\cap$$

$$\vdots$$

$$\cap$$

$$((S_m \cap S'_1) \rightarrow (T_m \cup T'_1)) \cap \dots \cap ((S_m \cap S'_{m'}) \rightarrow (T_m \cup T'_{m'}))$$

Which is a equivalent normalized function type:

$$((S_1 \cap S'_1) \rightarrow (T_1 \cup T'_1)) \cap \dots \cap ((S_{m \times m'} \cap S'_{m \times m'}) \rightarrow (T_{m \times m'} \cup T'_{m \times m'}))$$

as required.  $\square$

LEMMA 2.12. *For any function type there is equivalent normalized function type.*

PROOF. Cases:

- never: Is normalized function type.
- $S \rightarrow T$ : Is a normalized function type.
- $S \cap T$ : By indication we have normalized function types, and by definition normalized function types support intersection.
- $S \cup T$ : By indication we have normalized function types, and by Lemma 2.11 normalized function types support union.

This is an induction where the only tricky case is union, which is given by Lemma 2.11.  $\square$

Note that normalizing union can result in exponential blowup. We discuss some heuristics methods for avoiding space in §??, though we can't always avoid exponential blowup.

LEMMA 2.13. *For any type there is equivalent normalized type.*

$$\begin{aligned} \text{src}((S_1 \rightarrow T_1) \cap \dots \cap (S_m \rightarrow T_m)) &= S_1 \cup \dots \cup S_m \\ \text{app}((S_1 \rightarrow T_1) \cap \dots \cap (S_m \rightarrow T_m), V) &= U_1 \cap \dots \cap U_m \quad \text{where } U_i = \begin{cases} T_i & \text{if } V <: S_i \\ \text{unknown} & \text{otherwise} \end{cases} \end{aligned}$$

Fig. 6. The source  $\text{src}(F)$  and the application  $\text{app}(F, V)$  of a normalized function type  $F$  and argument  $V$ 

$$\begin{aligned} &(S_1 \rightarrow T_1) \cap \dots \cap (S_m \rightarrow T_m) \\ &\text{is saturated whenever} \\ &\forall i, j. \exists k. (S_i <: S_k) \wedge (S_j <: S_k) \wedge (T_i <: T_k) \wedge (T_j <: T_k) \end{aligned}$$

Fig. 7. A saturated normalized function type

PROOF. Cases:

- never: Is normalized type.
- unknown: Is equivalent normalized type  $(\text{never} \rightarrow \text{unknown}) \cup s_1 \cup \dots \cup s_n$ .
- $s$ : Is a normalized type  $\text{never} \cup s$ .
- $S \rightarrow T$ : Is a normalized type.
- $S \cap T$ : By indication  $S$  and  $T$ . We use Lemma 2.12 to normalize the function, and normalizes a union of scalar types is straightforward.
- $S \cup T$ : By indication  $S$  and  $T$ . We use Lemma 2.12 to normalize the function, and normalizes a interest of scalar types is straightforward..

This is an induction where the only tricky case is functions, which is given by Lemma 2.12.  $\square$

## 2.7 Saturated Normalization Function Types

The normalization function types, interact well with resolving overloads. For example:

$$F = (\text{number?} \rightarrow \text{number?}) \cap (\text{string?} \rightarrow \text{string?})$$

has two overloads. For example  $\text{app}$  in Fig 6 gives:

$$\begin{aligned} \text{app}(F, \text{number?}) &= \text{number?} \\ \text{app}(F, \text{string?}) &= \text{string?} \end{aligned}$$

This even works if more than resolving overload applies, for example:

$$\text{app}(F, \text{nil}) = \text{nil}$$

since:

$$\text{app}(F, \text{nil}) = (\text{number?} \cap \text{string?}) = \text{nil}$$

This works for intersection types, but unfortunately not union types. For example:

$$\text{app}(F, \text{number} \cup \text{string}) \text{ should be } \text{number?} \cup \text{string?}$$

but unfortunately  $\text{app}$  in Fig 6 will return unknown.

We consider the cases which for resolving overload the *saturated* normalized function types in Fig 7. For example  $F$  is not saturated there are overloads:

$$S_i = \text{number?} \quad S_j = \text{string?} \quad T_i = \text{number?} \quad T_j = \text{string?}$$

but there is no appropriate  $k$ . Fortunately, we can extend  $F$  to have an extra overload, to be saturated:

$$\begin{aligned} & (\text{number?} \rightarrow \text{number?}) \cap (\text{string?} \rightarrow \text{string?}) \\ & \cap ((\text{number?} \cup \text{string?}) \rightarrow (\text{number?} \cup \text{string?})) \end{aligned}$$

for which  $k$  is chosen:

$$S_k = (\text{number?} \cup \text{string?}) \quad T_k = (\text{number?} \cup \text{string?})$$

The saturated equivalent has the right resolving overload:

$$\text{app}(F, \text{number} \cup \text{string}) = \text{number?} \cup \text{string?}$$

There are two important results about saturated types. The first important result is every normalized function type can be converted to an equivalent saturated type.

LEMMA 2.14. *For any normalized function type there is equivalent saturated type.*

PROOF. Assume normalized function type:

$$F = (S_1 \rightarrow T_1) \cap \dots \cap (S_m \rightarrow T_m)$$

Let  $n$  be  $2^m$  as the powerset of  $1..m$ , as in Appendix A. We fold over the powerset of:

$$(S_{i+j} \rightarrow T_{i+j}) = ((S_i \cup S_j) \rightarrow (T_i \cup T_j))$$

To check that is saturated, for  $i$  and  $j$ , then pick  $k = i \boxplus j$  from which is direct that:

$$(S_i <: S_k) \wedge (S_j <: S_k) \wedge (T_i <: T_k) \wedge (T_j <: T_k)$$

To get the extension is equivalent, we use ??:

$$((S_1 \cup S_2) \rightarrow (T_1 \cup T_2)) <: ((S_1 \rightarrow T_1) \cap (S_2 \rightarrow T_2))$$

as required.  $\square$

The second important result is a test for  $F <: (S \rightarrow T)$  whenever  $F$  is saturated.

LEMMA 2.15. *For any saturated  $F$ ,  $S <: \text{src}(F)$  and  $\text{app}(F, S) <: T$  if and only if  $F <: (S \rightarrow T)$ .*

PROOF. Since  $F$  is saturated, there is an overload  $F <: (S_i \rightarrow T_i)$  where  $S <: S_i$  and  $T_i <: T$ .

( $\Rightarrow$ ) We hypothesize  $F <: (S \rightarrow T)$ . Lemma 2.8  $S <: S_i$  and  $T_i <: T$  and so by transitivity  $S <: \text{src}(F)$  and  $\text{app}(F, S) <: T$

( $\Leftarrow$ ) We hypothesize  $S <: \text{src}(F)$  and  $\text{app}(F) <: T$ . By Lemma 2.7  $(S_i \rightarrow T_i) <: (S \rightarrow T)$  and so by transitivity  $F <: (S \rightarrow T)$ .

Note Lemma 2.8 depends on pragmatic semantic subtyping.  $\square$

## 2.8 Algorithm for Deciding Subtypes of Normalization Types

We now have an algorithm for deciding subtyping.

THEOREM 2.16. *We have decision for subtyping.*

PROOF. We prove by induction on the depth of types, which means that  $(S \rightarrow T') <: (S' \rightarrow T)$  can use  $S <: S'$  and  $T <: T'$  by induction.

Cases:

- never  $<: T$ .
- $S \not<: \text{never}$  when  $S$  is a function or scalar.
- $S <: \text{unknown}$ .
- unknown  $\not<: T$  when  $T$  is a function or scalar.
- for scalars  $s$  and  $t$ , if  $s = t$  then  $s <: t$  otherwise  $s \not<: t$ .

- for scalar  $s$  and functions  $(S \rightarrow T)$ ,  $s \not\prec (S \rightarrow T)$  and  $(S \rightarrow T) \not\prec s$ .
- for functions  $(S_1 \rightarrow T_1)$  and  $(S_2 \rightarrow T_2)$ 
  - if  $S_2 \not\prec S_1$  then  $(S_1 \rightarrow T_1) \not\prec (S_2 \rightarrow T_2)$ ,
  - if  $T_1 \not\prec T_2$  then  $(S_1 \rightarrow T_1) \not\prec (S_2 \rightarrow T_2)$ , and
  - if  $S_2 \prec S_1$  and  $T_1 \prec T_2$  then  $(S_1 \rightarrow T_1) \prec (S_2 \rightarrow T_2)$ .
- comparing  $S$  and  $(T_1 \cap T_2)$ :
  - if  $S \not\prec T_1$  then  $S \not\prec (T_1 \cap T_2)$ ,
  - if  $S \not\prec T_2$  then  $S \not\prec (T_1 \cap T_2)$ , and
  - if  $S \prec T_1$  and  $S \prec T_2$  then  $S \prec (T_1 \cap T_2)$ .
- comparing  $(S_1 \cup S_2)$  and  $T$  as:
  - if  $S_1 \not\prec T$  then  $(S_1 \cup S_2) \not\prec T$ ,
  - if  $S_2 \not\prec T$  then  $(S_1 \cup S_2) \not\prec T$ , and
  - if  $S_1 \prec T$  and  $S_2 \prec T$  then  $(S_1 \cup S_2) \prec T$ .

Those are the cases, but this cases leaves where normalized cases, and saturated types for functions. The cases where normalization is used:

- comparing  $S$  and  $(T_1 \cup T_2)$ , and
- comparing  $(S_1 \cap S_2)$  and  $T$ .

Cases:

- comparing  $N$  and never:
  - if  $N$  is  $F \cup s_1 \cup \dots \cup s_n$ , then  $N \not\prec \text{never}$ , and
  - if  $N$  is never, then  $N \prec \text{never}$ .
- comparing unknown and  $N$ 
  - if  $N$  is never, then unknown  $\not\prec N$ ,
  - if  $N$  is  $F \cup s_1 \cup \dots \cup s_n$ , and there is missing  $s$  then unknown  $\not\prec N$ ,
  - if  $N$  is  $F \cup s_1 \cup \dots \cup s_n$ , and  $F \not\prec \text{function}$  then unknown  $\not\prec N$ , and
  - if  $N$  is  $F \cup s_1 \cup \dots \cup s_n$ , and  $F \prec \text{function}$  and ever scalar mentioned then unknown  $\prec N$ .
- comparing scalar  $s$  and  $N$ 
  - if  $N$  is never, then  $s \not\prec N$ ,
  - if  $N$  is  $F \cup s_1 \cup \dots \cup s_n$ , and there no  $s_i = s$ , then  $s \not\prec N$ , and
  - if  $N$  is  $F \cup s_1 \cup \dots \cup s_n$ , and there is  $s_i = s$ , then  $s \prec N$ .
- comparing  $N$  and  $(S \rightarrow T)$  we use the previous result.

as required. □

### 3 IMPLEMENTATION OF FULL LUAU

Stuff features

- Functions with arity.
- Functions with vararg.
- Tables.
- Tables with indexers.
- Tables with tagged unions.
- Singleton types.
- Generics.
- FFI.

Stuff imp's

- How we implement type normalization.
- Aim for type inference in 16ms.

- Telemetry with very few CodeTooComplex.
- SAT solves.

### 3.1 Function features: Oversaturation

Roblox offers a rich API for users to customize their experiences. Many of these APIs involve callbacks. A very simple example is that a developer can register a callback to be invoked when one 3D object comes into physical contact with another

```
part.Touched:Connect(function(other)
  print('This part was touched by ' .. other.Name)
end)
```

It is occasionally the case that this callback does not actually need all of the data passed to it. It is customary for the developer to simply omit the parameter:

```
part.Touched:Connect(function()
  print('This part was touched')
end)
```

The Roblox engine will at runtime oversaturate this function by passing it more arguments than it is designed to accept. This is perfectly safe to do but presents some challenges to type inference.

It is tempting to introduce a subtyping rule between functions of mismatched arities (eg  $() \rightarrow () <: T > (T) \rightarrow ()$ ), but this leads to soundness issues:

```
function print_string(x: string?)
  if x then
    print(string.lower(x))
  end
end
```

```
local a: () -> () = print_string
```

```
a(5) -- oversaturation is safe, right?
```

Instead, we observe that it is safe to oversaturate a value whose type is inferred from a function statement or expression. The only functions that are unsafe to oversaturate are the ones whose definitions we do not exactly know.

The simplest way to model a function that is safe to oversaturate is to consider it to be a variadic function that accepts any number of additional arguments of any types. We use the syntax *...unknown* to indicate 0 or more values of type *unknown*.

In the example above, we infer  $print\_string : (string?, \dots unknown) \rightarrow ()$ .

### 3.2 Function features: Undersaturation

In Luau, as in Lua, a function can also be called with fewer arguments than it has parameters. At runtime, this has same effect as passing *nil* for those extra parameters.

A very common practice is to use this mechanism to effect optional parameters.

```
function f(x)
  if ~= nil then
    print()
  end
end
```

```

687 f(0, true) -- Prints "true"
688 f(0, nil) -- Does not print anything
689 f(0)      -- Same as f(nil)

```

Luau considers an argument to be optional if its type is a subtype of *nil*. For instance, *f* above could be explicitly annotated like so:

```

693 function f<T>(x: T, y: boolean?): T

```

Modelling this in the type system is a challenge. Doing so via a subtyping relation  $\lambda a (T?) - > () <: () - > ()$  creates a soundness problem:

TODO remember what the issue was here!

### 3.3 Function features: variadic functions

We have already established that any function whose definition we exactly know is safe to oversaturate with any number of arguments of any types, and we have established that, via use of higher order functions and function aliasing, some functions are not safe to oversaturate, so it follows that there must be some subtyping relationship between variadic functions and fixed-arity functions. Without this, fixed-arity functions could not be used at all.

```

705 function apply_five(cb: (number) -> ())
706     cb(5)
707 end

```

```

709 apply_five(function() end)

```

Luau models functions as maps from one *type pack* to another. A type pack is an ordered set of 0 or more types. Like types, type packs can be concrete or generic. They can also be finite or infinite.

Subtype relations between finite type packs is simple and obvious: A type pack  $\bar{T}$  is a subtype of another  $\bar{U}$  if each type in  $\bar{T}$  is a subtype of the corresponding type from  $\bar{U}$ . There is no subtyping relation at all between finite packs of unequal length.

Function subtyping is standard except that we do a subtype test between type packs.

Luau does not directly support unions of type packs, but it is nevertheless very convenient to think about a variadic type pack as an infinite union of finite packs. If the pack  $\bar{T}$  is equal to  $\dots T$  (that is, 0 or more *T*s), then

$$\bar{T} = ()|(T)|(T, T)|(T, T, T)|\dots$$

And if we reapply the standard rule for subtyping of unions

$$\frac{\bar{T} <: \bar{A} \text{ or } \bar{T} <: \bar{B}}{\bar{T} <: \bar{A}|\bar{B}}$$

We can then continue to use the standard subtyping rule for functions:

$$\frac{\bar{C} <: \bar{A}, \bar{B} <: \bar{D}}{\bar{A} \rightarrow \bar{B} <: \bar{C} \rightarrow \bar{D}}$$

And we have an easy way to reason about tests like the following:

$$\begin{aligned}
 (A, B, \dots C) \rightarrow () <: (D, E) \rightarrow () \\
 (D, E) <: (A, B, \dots C) \\
 (D, E) <: (A, B)|(A, B, C)|(A, B, C, C)|(A, B, C, C, C)|\dots \\
 D <: A \text{ and } E <: B
 \end{aligned}$$

## 4 FURTHER WORK

### A POWERSET FOR FINITE SETS

A remap of standard exponenting on finite sets.

We assume a universe  $U$ . Let  $A \subseteq U$  be a finite set, where  $a_i \in U$ .

$$A = \{a_1, \dots, a_n\}$$

The exponential nonempty powerset  $\mathcal{P}(A) \subset \mathcal{P}(U)$ , where  $b_i \in \mathcal{P}(U)$ ;

$$\mathcal{P}(A) = \{b_1, \dots, b_{2^n}\}$$

We define  $b_i$  as:

- if  $i = 2^j$  then  $b_i = \{a_j\}$ , and
- if  $i = j + 2^k$  where  $j < 2^k$ , then  $b_i = b_j \cup \{a_k\}$ .

A common case is a fold over  $(c_i \oplus c_j)$  where  $c_i \in U$ ,  $c_j \in U$ , and  $c_i \oplus c_j \in U$ :

- if  $i = 2^j$  then  $b_i = a_j$ , and
- if  $i = j + 2^k$  where  $j < 2^k$ , then  $b_i = b_j \oplus a_k$ .

Another common case is where  $i$  and  $j$  exists, and find  $(i \boxplus j)$  where:

$$b_i \cup b_j = b_{i \boxplus j}$$

We define  $(i \boxplus j)$  as:

- if  $i = 0$  then  $(i \boxplus j)$  is  $j$ ,
- if  $j = 0$  then  $(i \boxplus j)$  is  $i$ ,
- if  $i = i' + 2^k$  and  $j = j' + 2^k$  then  $(i \boxplus j)$  is  $(i' \boxplus j') + 2^k$ ,
- if  $i = i' + 2^k$  and  $j < 2^k$  then  $(i \boxplus j)$  is  $(i' \boxplus j) + 2^k$ , and
- if  $i < 2^k$  and  $j = j' + 2^k$  then  $(i \boxplus j)$  is  $(i \boxplus j') + 2^k$ .

These allow us to unions of exponential nonempty powerset, while treat using indexes.

## REFERENCES

- [1] L. Brown, A. Friesen, and A. S. A. Jeffrey. 2023. Goals of the Luau Type System: Two Years Only. In *Proc. Human Aspects of Types and Reasoning Assistants*. <https://asaj.org/papers/hatra23.pdf>
- [2] L. Brown and A. S. A. Jeffrey. 2023. Luau Prototype Typechecker. <https://github.com/luau-lang/agda-typeck>
- [3] G. Castagna and A. Frisch. 2005. A Gentle Introduction to Semantic Subtyping. In *Proc. Principles and Practice of Declarative Programming*.
- [4] A. Frisch, G. Castagna, and V. Benzaken. 2008. Semantic subtyping: Dealing set-theoretically with function, union, intersection, and negation types. *J. ACM* 55, 19 (2008).
- [5] A. S. A. Jeffrey. 2022. Semantic Subtyping in Luau. Roblox Technical Blog. <https://blog.roblox.com/2022/11/semantic-subtyping-luau/>
- [6] A. M. Kent. 2021. Down and Dirty with Semantic Set-theoretic Types (a tutorial). <https://pnwamk.github.io/sst-tutorial/>
- [7] Roblox. 2022. Roblox Coporation 2022 Annual Report. <https://ir.roblox.com/financials/annual-reports/default.aspx>