Lecture 6: Set-theoretic Semantics, Categories

Lecturer: Max S. New
Scribe: Jonah Nan

Jan 30, 2023

Our main goal for this lecture is to provide a set-theoretic *semantics* for Simple Type Theory (STT) and to prove that this semantic interpretation is *sound*. As a corollary, we obtain the *consistency* of STT's theory of equality of terms.

We will define the set-theoretic semantics of STT in several stages. First, we inductively define the denotation of a type. The denotation of a type will be a set. To give a denotation to base types $X \in \Sigma_0$, we require an assignment $\sigma_0 : \Sigma_0 \to \text{Set}$ which maps each base type to some set. We define:

$$\begin{aligned}
[\![1]\!] &:= \{*\} && \text{(The singleton set)}\\
[\![0]\!] &:= \emptyset && \text{(The empty set)}\\
[\![A \times B]\!] &:= [\![A]\!] \times [\![B]\!] && \text{(The Cartesian product)}\\
[\![A + B]\!] &:= [\![A]\!] \uplus [\![B]\!] = (\{1\} \times [\![A]\!]) \cup (\{2\} \times [\![B]\!]) && \text{(The disjoint union)}\\
[\![A \Rightarrow B]\!] &:= [\![B]\!]^{[\![A]\!]} && \text{(The set of functions from } [\![A]\!] \text{ to } [\![B]\!])\\
[\![X]\!] &:= \sigma_0(X) && \text{(For X a base type)}
\end{aligned}$$

Next, we define the denotation of a context Γ. Let $\Gamma = x_1 : A_1, ..., x_n : A_n$. We define:

$$[\![\Gamma]\!] := \prod_{i=1}^{n} [\![A_i]\!]$$

As a special case, $[\![\cdot]\!] = \{*\}$, as the empty product of sets is the singleton set. An element $\tilde{\gamma} \in [\![\Gamma]\!]$ is a tuple of n elements whose i-th element belongs to the set $[\![A_i]\!]$. We can think of $\tilde{\gamma}$ as an assignment of variables so that each x_i is mapped to a member of the corresponding set $[\![A_i]\!]$. We use the notation $\tilde{\gamma}(x_i) \in [\![A_i]\!]$ to refer to the i-th member of the tuple $\tilde{\gamma}$. We refer to $\tilde{\gamma}$ as a *semantic substitution*, in contrast to the *syntactic substitution* $\gamma : \Delta \to \Gamma$ defined in PS2.

We now define the denotation of a term M, again by induction. For any judgement $\Gamma \vdash M : A$, we will give a corresponding denotation $[\![M]\!] : [\![\Gamma]\!] \to [\![A]\!]$. (Formally we should write this as $[\![M]\!]^{\Gamma}$, but these are usually clear from context). Since $[\![M]\!]$ will be a function operating on semantic substitutions $\tilde{\gamma} \in [\![\Gamma]\!]$, it suffices to define the action of $[\![M]\!]$ on each $\tilde{\gamma}$.

For $\Gamma \vdash x : A$, we define

$$[\![x]\!](\tilde{\gamma}) := \tilde{\gamma}(x)$$

By assumption $x : A \in \Gamma$, and hence $\tilde{\gamma}(x) \in [\![A]\!]$. To give a denotation to terms involving application of a function symbol, we require an assignment σ_1 which maps each function symbol $(f : A_1, ..., A_n \to B) \in \Sigma_1$ to a function $\sigma_1(f) : (\prod_{i=1}^n [\![A_i]\!]) \to [\![B]\!]$. For $\Gamma \vdash f(M_1, ..., M_n) : B$, we define

$$[\![f(M_1, ..., M_n)]\!](\tilde{\gamma}) := (\sigma_1(f))([\![M_1]\!](\tilde{\gamma}), ..., [\![M_n]\!](\tilde{\gamma}))$$

By assumption each $[\![M_i]\!](\tilde{\gamma}) \in A_i$, and so we can apply the function $\sigma_1(f) : (\prod_{i=1}^n [\![A_i]\!]) \to [\![B]\!]$ to obtain a member of $[\![B]\!]$.

For $\Gamma \vdash () : 1$, we define

$$[\![()]\!](\tilde{\gamma}) := * \in \{*\} = [\![1]\!]$$

For $\Gamma \vdash \text{case}_0 M\{\} : A$, we claim that $[\![\Gamma]\!] = \emptyset$, and hence we define

$$[\![\text{case}_0 M\{\}]\!] := (* : \emptyset \to [\![A]\!])$$

This may seem strange, but recall that in set theory there is exactly one function from $\emptyset$ to any given set S; the set of functions from $\emptyset$ to S is $S^\emptyset = \{*\}$, the singleton set. So we just need to show that $[\![\Gamma]\!] = \emptyset$. By assumption, we have $[\![M]\!] : [\![\Gamma]\!] \to [\![0]\!]$, so $[\![M]\!] : [\![\Gamma]\!] \to \emptyset$. If S is a nonempty set, then there exist no functions from S to $\emptyset$, so we must have $[\![\Gamma]\!] = \emptyset$.

For $\Gamma \vdash (M_1, M_2) : A_1 \times A_2$, we define

$$[\![(M_1, M_2)]\!](\tilde{\gamma}) := ([\![M_1]\!](\tilde{\gamma}), [\![M_2]\!](\tilde{\gamma}))$$

By assumption, $[\![M_1]\!] : [\![\Gamma]\!] \to [\![A_1]\!]$ and $[\![M_2]\!] : [\![\Gamma]\!] \to [\![A_2]\!]$. Hence $([\![M_1]\!](\tilde{\gamma}), [\![M_2]\!](\tilde{\gamma})) \in [\![A_1]\!] \times [\![A_2]\!] = [\![A_1 \times A_2]\!]$.

For $\Gamma \vdash \pi_i N : A_i,\ i \in \{1, 2\}$, we define

$$[\![\pi_i N]\!](\tilde{\gamma}) := \pi_i([\![N]\!](\tilde{\gamma}))$$

By assumption, $[\![N]\!] : [\![\Gamma]\!] \to ([\![A_1 \times A_2]\!] = [\![A_1]\!] \times [\![A_2]\!])$, so the output $[\![N]\!](\tilde{\gamma})$ is an ordered pair whose i-th component lies in $[\![A_i]\!]$.

For $\Gamma \vdash \lambda x.M : A \Rightarrow B$, we wish to give $[\![\lambda x.M]\!](\tilde{\gamma}) \in ([\![A \Rightarrow B]\!] = [\![B]\!]^{[\![A]\!]})$. So we should have $[\![\lambda x.M]\!](\tilde{\gamma}) : [\![A]\!] \to [\![B]\!]$. It thus suffices to define how $[\![\lambda x.M]\!](\tilde{\gamma})$ acts on each $\tilde{x} \in [\![A]\!]$. We define

$$([\![\lambda x.M]\!](\tilde{\gamma}))(\tilde{x}) := [\![M]\!](\tilde{\gamma}, \tilde{x}/x)$$

Here the notation $(\tilde{\gamma}, \tilde{x}/x)$ means extending the assignment $\tilde{\gamma}$ to take one additional input, x, and map it to $\tilde{x} \in [\![A]\!]$. Thus we have $(\tilde{\gamma}, \tilde{x}/x) \in [\![\Gamma]\!] \times [\![A]\!] = [\![\Gamma, x : A]\!]$. By assumption, $[\![M]\!] : [\![\Gamma, x : A]\!] \to [\![B]\!]$, and hence $[\![M]\!](\tilde{\gamma}, \tilde{x}/x) \in [\![B]\!]$, as desired.

For $\Gamma \vdash M\, N : B$, we define

$$[\![MN]\!](\tilde{\gamma}) := ([\![M]\!](\tilde{\gamma}))([\![N]\!](\tilde{\gamma}))$$

By assumption, we have $[\![N]\!] : [\![\Gamma]\!] \to [\![A]\!]$, so $[\![N]\!](\tilde{\gamma}) \in [\![A]\!]$. Also by assumption, we have

$[\![M]\!] : [\![\Gamma]\!] \to ([\![A \Rightarrow B]\!] = [\![B]\!]^{[\![A]\!]})$. Thus we have $[\![M]\!](\tilde{\gamma}) : [\![A]\!] \to [\![B]\!]$. Thus $([\![M]\!](\tilde{\gamma}))([\![N]\!](\tilde{\gamma})) \in [\![B]\!]$, as desired.

For $\Gamma \vdash i_j M_j : A_1 + A_2$, $j \in \{1, 2\}$, we define

$$[\![i_j M_j]\!](\tilde{\gamma}) := (j, [\![M_j]\!](\tilde{\gamma}))$$

By assumption, $[\![M_j]\!] : [\![\Gamma]\!] \to [\![A_j]\!]$, so $(j, [\![M_j]\!](\tilde{\gamma})) \in \{j\} \times [\![A_j]\!] \subseteq (\{1\} \times [\![A_1]\!]) \cup (\{2\} \times [\![A_2]\!]) = [\![A_1]\!] \uplus [\![A_2]\!] = [\![A_1 + A_2]\!]$.

For $\Gamma \vdash \text{case}_+ M\{i_1 x_1 \to N_1 | i_2 x_2 \to N_2\} : B$, we define

$$[\![\text{case}_+ M\{i_1 x_1 \to N_1 | i_2 x_2 \to N_2\}]\!](\tilde{\gamma}) := \begin{cases} [\![N_1]\!](\tilde{\gamma}, \pi_2([\![M]\!](\tilde{\gamma}))/x_1), & \pi_1([\![M]\!](\tilde{\gamma})) = 1 \\ [\![N_2]\!](\tilde{\gamma}, \pi_2([\![M]\!](\tilde{\gamma}))/x_2), & \pi_1([\![M]\!](\tilde{\gamma})) = 2 \end{cases}$$

By assumption, $[\![M]\!] : [\![\Gamma]\!] \to ([\![A_1 + A_2]\!] = [\![A_1]\!] \uplus [\![A_2]\!])$. So $[\![M]\!](\tilde{\gamma}) \in [\![A_1]\!] \uplus [\![A_2]\!]$. This means $\pi_1([\![M]\!](\tilde{\gamma}))$ is either 1 or 2. Let $j = \pi_1([\![M]\!](\tilde{\gamma})) \in \{1, 2\}$. Then $\pi_2([\![M]\!](\tilde{\gamma})) \in [\![A_j]\!]$. So $(\tilde{\gamma}, \pi_2([\![M]\!](\tilde{\gamma}))/x_j) \in [\![\Gamma]\!] \times [\![A_j]\!] = [\![\Gamma, x_j : A_j]\!]$. By assumption, $[\![N_j]\!] : [\![\Gamma, x_j : A_j]\!] \to [\![B]\!]$. So $[\![N_j]\!](\tilde{\gamma}, \pi_2([\![M]\!](\tilde{\gamma}))/x_j) \in [\![B]\!]$, as desired.

This completes the inductive definition of denotation of terms $[\![M]\!]$:

Theorem 1 (Well-definedness of denotations)**.** *If* $\Gamma \vdash M : A$ *then* $[\![M]\!]^\Gamma : [\![\Gamma]\!] \to [\![A]\!]$

As a corollary of this, we obtain one part of the consistency of STT as a logic:

Corollary 1. *There is no* M *such that* $\cdot \vdash M : 0$.

Proof. If $\cdot \vdash M : 0$, then $[\![M]\!] : [\![\cdot]\!] \to [\![0]\!]$, i.e., $[\![M]\!] : \{*\} \to \emptyset$, but there is no such function. ☐

Our next step will be to prove the *compositionality* theorem, which will be an important lemma used in our proof of soundness. This will rely on the notion of syntactic substitution defined in PS2.

For a syntactic substitution $\gamma : \Delta \to \Gamma$, we define the denotation $[\![\gamma]\!] : [\![\Delta]\!] \to [\![\Gamma]\!]$. $[\![\gamma]\!]$ takes a semantic substitution $\tilde{\delta} \in [\![\Delta]\!]$ and maps it to a semantic substitution $[\![\gamma]\!](\tilde{\delta}) \in [\![\Gamma]\!]$. To define the output $[\![\gamma]\!](\tilde{\delta}) \in [\![\Gamma]\!]$, we define how it maps each $x_i : A_i \in \Gamma$ to a member of $[\![A_i]\!]$. We define:

$$([\![\gamma]\!](\tilde{\delta}))(x_i) := [\![\gamma(x_i)]\!](\tilde{\delta})$$

By the definition of syntactic substitution, for each $x_i : A_i \in \Gamma$ we have $\Delta \vdash \gamma(x) : A_i$, so we have $[\![\gamma(x)]\!] : [\![\Delta]\!] \to [\![A_i]\!]$. Thus $[\![\gamma(x)]\!](\tilde{\delta}) \in [\![A_i]\!]$, as desired.

We can now state the compositionality theorem. We know from PS2 that whenever $\Gamma \vdash M : B$, we have $\Delta \vdash M[\gamma] : B$. We thus have $[\![\gamma]\!] : [\![\Delta]\!] \to [\![\Gamma]\!]$, $[\![M]\!] : [\![\Gamma]\!] \to [\![B]\!]$, and $[\![M[\gamma]]\!] : [\![\Delta]\!] \to [\![B]\!]$.

Lemma 1 (Compositionality of the Set-theoretic Semantics)**.** *If* $\Gamma \vdash M : A$ *and* $\gamma : \Delta \to \Gamma$*, then*

$$[\![M[\gamma]]\!]^\Delta = [\![M]\!]^\Gamma \circ [\![\gamma]\!]^\Delta$$

I.e., for any $\tilde{\delta} \in [\![\Delta]\!]$, $[\![M[\gamma]]\!](\tilde{\delta}) = [\![M]\!]([\![\gamma]\!](\tilde{\delta}))$.

As a special case, when $\Gamma \vdash N : A$ and $\Gamma, x : A \vdash M : B$ then we have the substitution $(\mathrm{id}_\Gamma, N/x) : \Gamma \to \Gamma, x : A$. We have $M[N/x] = M[(\mathrm{id}_\Gamma, N/x)]$. Note that $[\![(\mathrm{id}_\Gamma, N/x)]\!](\tilde{\gamma}) = (\tilde{\gamma}, [\![N]\!](\tilde{\gamma})/x)$. (you can check this). Thus, applying compositionality here gives:

$$[\![M[N/x]]\!](\tilde{\gamma}) = [\![M]\!](\tilde{\gamma}, [\![N]\!](\tilde{\gamma})/x)$$

We prove compositionality by induction on M.

- If $M = x$ is a variable:

$$\begin{aligned}
[\![x[\gamma]]\!](\tilde{\delta}) &= [\![\gamma(x)]\!](\tilde{\delta}) \\
&= ([\![\gamma]\!](\tilde{\delta}))(x) \\
&= [\![x]\!]([\![\gamma]\!](\tilde{\delta}))
\end{aligned}$$

- If $M = f(M_1, ..., M_n)$ is an application of a function symbol:

$$\begin{aligned}
[\![f(M_1, ..., M_n)[\gamma]]\!](\tilde{\delta}) &= [\![f(M_1[\gamma], ..., M_n[\gamma])]\!](\tilde{\delta}) \\
&= (\sigma_1(f))([\![M_1[\gamma]]\!](\tilde{\delta}), ..., [\![M_n[\gamma]]\!](\tilde{\delta})) \\
&= (\sigma_1(f))([\![M_1]\!]([\![\gamma]\!](\tilde{\delta})), ..., [\![M_n]\!]([\![\gamma]\!](\tilde{\delta}))) \\
&\qquad \text{(by inductive hypothesis)} \\
&= [\![f(M_1, ..., M_n)]\!]([\![\gamma]\!](\tilde{\delta}))
\end{aligned}$$

- If $M = ()$:

$$\begin{aligned}
[\![()[\gamma]]\!](\tilde{\delta}) &= [\![()]\!](\tilde{\delta}) \\
&= * \\
&= [\![()]\!]([\![\gamma]\!](\tilde{\delta}))
\end{aligned}$$

- If $M = \mathrm{case}_0 N\{\}$: As above, we know $[\![\Delta]\!] = \emptyset$, and any two functions $f, g : \emptyset \to [\![B]\!]$ are equal. (We have $f = * = g$.) We have both $[\![M[\gamma]]\!] : \emptyset \to [\![B]\!]$ and $[\![M]\!] \circ [\![\gamma]\!] : \emptyset \to [\![B]\!]$, so the result follows.

- If $M = (M_1, M_2)$ is a pair:

$$\begin{aligned}
[\![(M_1, M_2)[\gamma]]\!](\tilde{\delta}) &= [\![(M_1[\gamma], M_2[\gamma])]\!](\tilde{\delta}) \\
&= ([\![M_1[\gamma]]\!](\tilde{\delta}), [\![M_2[\gamma]]\!](\tilde{\delta})) \\
&= ([\![M_1]\!]([\![\gamma]\!](\tilde{\delta})), [\![M_2]\!]([\![\gamma]\!](\tilde{\delta}))) && \text{(ind. hyp.)} \\
&= [\![(M_1, M_2)]\!]([\![\gamma]\!](\tilde{\delta}))
\end{aligned}$$

- If $M = \pi_i N$ is a projection:

$$\begin{aligned}
[\![(\pi_i N)[\gamma]]\!](\tilde{\delta}) &= [\![\pi_i(N[\gamma])]\!](\tilde{\delta}) \\
&= \pi_i([\![N[\gamma]]\!](\tilde{\delta})) \\
&= \pi_i([\![N]\!]([\![\gamma]\!](\tilde{\delta}))) && \text{(ind. hyp.)} \\
&= [\![\pi_i N]\!]([\![\gamma]\!](\tilde{\delta}))
\end{aligned}$$

- The case of $\lambda x.M$ is interesting. This case is the reason we use multi-variable substitutions γ in the statement of the theorem: the weaker inductive statement with single-variable substitutions is less straightforward to prove in this case.

  We have $[\![(\lambda x.M)[\gamma]]\!](\tilde{\delta}) = [\![\lambda x.(M[\gamma, x/x])]\!](\tilde{\delta})$. This and $[\![\lambda x.M]\!]([\![\gamma]\!]\tilde{\delta})$ are both functions $[\![A]\!] \to [\![B]\!]$, so it suffices to show they act the same on each $\tilde{x} \in [\![A]\!]$.

  $$\begin{aligned}([\![\lambda x.(M[\gamma, x/x])]\!](\tilde{\delta}))(\tilde{x}) &= [\![M[\gamma, x/x]]\!](\tilde{\delta}, \tilde{x}/x) \\ &= [\![M]\!]([\![\gamma, x/x]\!](\tilde{\delta}, \tilde{x}/x)) && \text{(ind. hyp.)}\end{aligned}$$

  On the other hand, $([\![\lambda x.M]\!]([\![\gamma]\!]\tilde{\delta}))(\tilde{x}) = [\![M]\!]([\![\gamma]\!]\tilde{\delta}, \tilde{x}/x)$. So we need only show $[\![\gamma, x/x]\!](\tilde{\delta}, \tilde{x}/x) = ([\![\gamma]\!]\tilde{\delta}, \tilde{x}/x)$. To do this, we show they act the same on each $y_i : C_i \in (\Gamma, x : A)$. We have:

  $$\begin{aligned}([\![\gamma, x/x]\!](\tilde{\delta}, \tilde{x}/x))(x) &= [\![(\gamma, x/x)(x)]\!](\tilde{\delta}, \tilde{x}/x) \\ &= [\![x]\!](\tilde{\delta}, \tilde{x}/x) \\ &= (\tilde{\delta}, \tilde{x}/x)(x) \\ &= \tilde{x}\end{aligned}$$

  Likewise, $([\![\gamma]\!]\tilde{\delta}, \tilde{x}/x)(x) = \tilde{x}$. Finally, for any $y_i : C_i \in \Gamma$ (so $y_i \neq x$), we have $([\![\gamma, x/x]\!](\tilde{\delta}, \tilde{x}/x))(y_i) = [\![(\gamma, x/x)(y_i)]\!](\tilde{\delta}, \tilde{x}/x) = [\![\gamma(y_i)]\!](\tilde{\delta}, \tilde{x}/x)$. On the other hand, $([\![\gamma]\!]\tilde{\delta}, \tilde{x}/x)(y_i) = ([\![\gamma]\!]\tilde{\delta})(y_i) = [\![\gamma(y_i)]\!](\tilde{\delta})$.

  So we need to show that $[\![\gamma(y_i)]\!](\tilde{\delta}, \tilde{x}/x) = [\![\gamma(y_i)]\!](\tilde{\delta})$. Here it is useful to use the annotations, what we need to show is:

  $$[\![\gamma(y_i)]\!]^{\Delta, x:A}(\tilde{\delta}, \tilde{x}/x) = [\![\gamma(y_i)]\!]^{\Delta}(\tilde{\delta})$$

  where $\gamma : \Delta \to \Gamma$ and $y : B \in \Gamma$. So on the left hand side, $\gamma(y)$ is implicitly weakened, but they are interpreted as functions with different domains. Intuitively equation is valid since $\gamma(y)$ doesn't use the variable x. Formally, we need to prove it as a lemma.

- The other cases are similar to the previous. Most complex is the + elimination case, which is similar to λ since it involves variable binding.

Lemma 2 (Weakening). *If Δ' contains every variable in Δ, and $\Delta' \vdash M : A$, then*

$$[\![M]\!]^{\Delta'}\tilde{\delta}' = [\![M]\!]^{\Delta}(\tilde{\delta}'|_{\Delta})$$

where $\tilde{\delta}'|_{\Delta}$ is the restriction of the tuple $\tilde{\delta}'$ to only the fields which are variables in Δ.

Proof. By induction on M. Similar to the compositionality proof. Most interesting is the λ case: If $M = \lambda x.N$, then

$$\begin{aligned}&[\![\lambda x.N]\!]^{\Delta'}(\tilde{\delta}')(\tilde{x}) \\ &\quad = [\![N]\!]^{\Delta', x:A}(\tilde{\delta}', \tilde{x}/x) \\ &\quad = [\![N]\!]^{\Delta, x:A}(\tilde{\delta}'|_{\Delta}, \tilde{x}/x) && \text{(ind. hyp.)} \\ &\quad = [\![\lambda x.N]\!]^{\Delta, x:A}(\tilde{\delta}'|_{\Delta})(\tilde{x})\end{aligned}$$

□

Armed with the compositionality theorem, we're ready to prove the soundness of our model of the equational theory of STT.

Theorem 2 (Soundness of Equational Theory for Set-theoretic Semantics)**.** *Suppose that for all* $(\Gamma, M, N, A) \in \Sigma_2$ *we have* $[\![M]\!] = [\![N]\!]$*. Then whenever* $\Gamma \vdash M = N : A$*, we have* $[\![M]\!] = [\![N]\!]$*.*

Proof. As usual, we proceed by induction on the proof that $\Gamma \vdash M = N : A$.

- The cases of the reflexive, symmetric, and transitive deduction rules follow immediately from the reflexive, symmetric, and transitive properties of set-theoretic equality.
- The case of axioms $\in \Sigma_2$ holds by assumption.
- The cases of the congruence rules all follow from the substitution property of set-theoretic equality. (Note that SubstCong requires an application of compositionality.) For example, the case of $\times$ I Cong:

  Suppose $(M, N) = (M', N')$ by $\times$ I Cong. We know that $M = M'$ and $N = N'$. By inductive hypothesis, $[\![M]\!] = [\![M']\!]$ and $[\![N]\!] = [\![N']\!]$. So

  $$\begin{aligned} [\![(M, N)]\!](\tilde{\gamma}) &= ([\![M]\!](\tilde{\gamma}), [\![N]\!](\tilde{\gamma})) \\ &= ([\![M']\!](\tilde{\gamma}), [\![N']\!](\tilde{\gamma})) \\ &= [\![(M', N')]\!](\tilde{\gamma}) \end{aligned}$$

  hence $[\![(M, N)]\!] = [\![(M', N')]\!]$.
- The case of $\Rightarrow \beta$: We know $\Gamma, x : A \vdash M : B$ and $\Gamma \vdash N : A$. We conclude $(\lambda x.M)\ N = M[N/x]$. We have: $[\![(\lambda x.M)\ N]\!](\tilde{\gamma}) = ([\![\lambda x.M]\!](\tilde{\gamma}))([\![N]\!](\tilde{\gamma})) = [\![M]\!](\tilde{\gamma}, [\![N]\!](\tilde{\gamma})/x)$ By compositionality, this $= [\![M[N/x]]\!](\tilde{\gamma})$.
- The case of $\Rightarrow \eta$: We know $\Gamma \vdash M : A \Rightarrow B$. We conclude $\Gamma \vdash M = \lambda x.(M\ x) : A \Rightarrow B$. We have: $[\![\lambda x.(M\ x)]\!](\tilde{\gamma}) = (\tilde{x} \mapsto [\![M\ x]\!](\tilde{\gamma}, \tilde{x}/x)) = (\tilde{x} \mapsto ([\![M]\!](\tilde{\gamma}, \tilde{x}/x))([\![x]\!](\tilde{\gamma}, \tilde{x}/x))) = (\tilde{x} \mapsto ([\![M]\!](\tilde{\gamma}, \tilde{x}/x))(\tilde{x})) = (\tilde{x} \mapsto ([\![M]\!](\tilde{\gamma}))(\tilde{x})) = [\![M]\!](\tilde{\gamma})$.

  TODO: I think this is another application of that same lemma.
- The remaining β and η rules are left as an exercise.

□

As a corollary of soundness, we see it is *impossible* to prove $i_1() = i_2()$ in STT, hence the theory of equality is consistent.

Corollary 2. $\cdot \vdash i_1() = i_2() : 1 + 1$ *is not provable in STT with no axioms.*

Proof. If $i_1() = i_2()$ is provable, then by the soundness of the equational theory $[\![i_1()]\!] = [\![i_2()]\!]$, but
$$[\![i_1()]\!](*) = (1, *) \neq (2, *) = [\![i_2()]\!](*)$$
□

But there are more models of STT beyond this intuitive set-theoretic one. To describe them, we need to learn... Category Theory!

Definition 1. *A category $\mathcal{C}$ consists of:*

1. *$\mathcal{C}_0$, a set of objects*
2. *For each $a, b \in \mathcal{C}_0$, a set $\mathcal{C}_1(a, b)$ of arrows (aka morphisms) from a to b. For $f \in \mathcal{C}_1(a, b)$ when the category is clear from context, we write $f : a \to b$.*
3. *For each $a \in \mathcal{C}_0$ a distinguished identity morphism $id_a \in \mathcal{C}_1(a, a)$.*
4. *For each $a, b, c \in \mathcal{C}_0$, a composition operation $\circ : (\mathcal{C}_1(b, c) \times \mathcal{C}_1(a, b)) \to \mathcal{C}_1(a, c)$*
5. *Composition respects the identity morphisms: for any $f : a \to b$, we have*
$$id_b \circ f = f$$
and
$$f \circ id_a = f$$
6. *Composition is associative: wherever the composition is defined, we have $f \circ (g \circ h) = (f \circ g) \circ h$.*

There are foundational issues with formalizing category theory in terms of set theory. We wish to have a "category of all sets", but then its set of objects would need to be a set containing all sets... this is problematic. Instead, we consider a category of all "small" sets, and this is good enough for any practical purposes. There are a lot of neat foundational things happening here, but they won't be focused on in this course.

We can view any preorder $(X, \leq)$ as a category. The objects in our category are the elements of the preorder's underlying set, and we have a single morphism $* : a \to b$ exactly when $a \leq b$. The reflexivity of $\leq$ ensures the existence of identity morphisms, and the transitivity of $\leq$ ensures that we can define a composition operation. The fact that composition respects the identity morphisms and is associative is clear because for any given source and target there is at most one possible morphism, so any two morphisms with the same source and target must be equal.

In this way, category theory generalizes order theory in the same way that simple type theory generalizes IPL. In order theory we only care *if* $x \leq y$ holds, but in category theory we care about which morphism we have in $\mathcal{C}_1(a, b)$. Similarly in IPL we only care *if* $\Gamma \vdash A$ is provable, but in STT we care about which program we have $\Gamma \vdash M : A$