Type Safety

Concepts of Programming Languages
Lecture 19

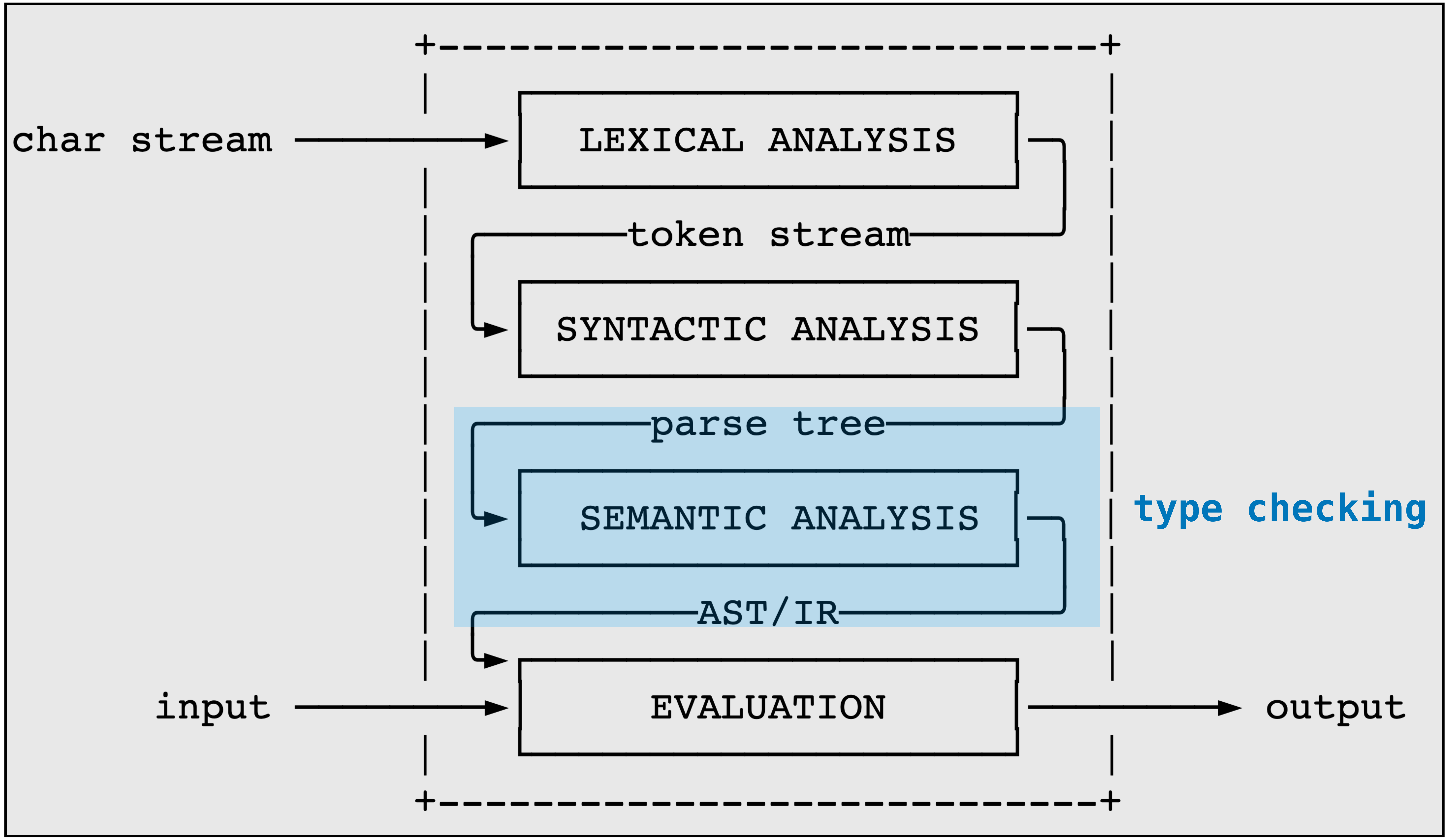
CAS CS 320

Outline

- Demo an **implementation** of the simply typed lambda calculus and desugaring
- Discuss **induction** over derivations
- Show that STLC satisfies **progress** and **preservation**

Recap

Recall: The Picture



Recall: Type Checking/Inference

```
type_check : expr -> ty -> bool

type_of : expr -> ty option
```

Recall: Type Checking/Inference

```
type_check : expr -> ty -> bool

type_of : expr -> ty option
```

<u>Type checking:</u> determining whether an expression can be typed

Recall: Type Checking/Inference

```
type_check : expr -> ty -> bool

type_of : expr -> ty option
```

<u>Type checking:</u> determining whether an expression can be typed

<u>Type inference:</u> *synthesizing* a type for an expression

Recall: Type Checking/Inference

```
type_check : expr -> ty -> bool

type_of : expr -> ty option
```

Type checking: determining whether an expression can be typed

Type inference: *synthesizing* a type for an expression

Theoretically, these two problems can be very different. *For STLC, they are both easy*

Recall: Syntax (STLC)

fun (x : τ) → e

$$e ::= \bullet \mid x \mid \lambda x^{\tau}.\, e \mid ee$$

$$\tau ::= \top \mid \tau \to \tau$$

$$x ::= \textit{variables}$$

The syntax is the same as that of the lambda calculus except:

» we include a unit expression

» we have types, which annotate arguments

Recall: Typing (STLC)

$$\frac{}{\Gamma \vdash \bullet : \top}\ \textbf{unit}$$

$$\frac{\Gamma, x : \tau \vdash e : \tau'}{\Gamma \vdash \lambda x^{\tau}. e : \tau \to \tau'}\ \textbf{abstraction}$$

$$\frac{(x : \tau) \in \Gamma}{\Gamma \vdash x : \tau}\ \textbf{variable}$$

$$\frac{\Gamma \vdash e_1 : \tau \to \tau' \qquad \Gamma \vdash e_2 : \tau}{\Gamma \vdash e_1 e_2 : \tau'}\ \textbf{application}$$

These rules enforce that a function can only be applied if we *know* that it's a function

Recall: Semantics (STLC)

$$\frac{e_1 \longrightarrow e_1'}{e_1 e_2 \longrightarrow e_1' e_2}\ \texttt{leftEval}$$

$$\frac{}{(\lambda x \,.\, e)e' \longrightarrow [e'/x]e}\ \texttt{beta}$$

Recall: Semantics (STLC)

$$\frac{e_1 \longrightarrow e_1'}{e_1 e_2 \longrightarrow e_1' e_2}\ \texttt{leftEval} \qquad \frac{}{(\lambda x^{\tau}.\, e)e' \longrightarrow [e'/x]e}\ \texttt{beta}$$

`We can use any semantics (this is small-step CBN)`

Recall: Semantics (STLC)

$$\frac{e_1 \longrightarrow e_1'}{e_1 e_2 \longrightarrow e_1' e_2}\ \texttt{leftEval} \qquad \frac{}{(\lambda x^{\tau}.\, e)e' \longrightarrow [e'/x]e}\ \texttt{beta}$$

We can use any semantics (this is small-step CBN)

This is part of the point. Type-checking only determines *whether* we go on to evaluate the program

Recall: Semantics (STLC)

$$\frac{e_1 \longrightarrow e_1'}{e_1 e_2 \longrightarrow e_1' e_2}\ \texttt{leftEval}$$

$$\frac{}{(\lambda x^{\tau}.\, e)e' \longrightarrow [e'/x]e}\ \texttt{beta}$$

We can use any semantics (this is small-step CBN)

This is part of the point. Type-checking only determines *whether* we go on to evaluate the program

It doesn't determine *how* we evaluate the program

Practice Problem

$$(\lambda x^{(\top \to \top) \to \top} . x(\lambda z^{\top} . x(wz)))y$$

Determine the smallest context such that the above expression is typeable (also give its type)

Answer

$$(\lambda x^{(\top \to \top)\to \top} . x(\lambda z^{\top} . x(wz)))\,y$$

$\top$

$(\top \to \top) \to \top$

$$\frac{\Gamma \vdash e_1 : \tau \to \tau' \qquad \Gamma \vdash e_2 : \tau}{\Gamma \vdash e_1 e_2 : \tau'}$$

$\top \to \top$

$w : ?,\ x : (\top \to \top) \to \top \vdash x\,(\lambda z^{\top} . \,x\,(w\,z)) : ??$

$\top \to \top$

$\top$

do the derivation

$\top \to \top \to \top$

$\{y : (\top \to \top) \to \top,\ w : \top \to (\top \to \top)\} \vdash (\lambda x^{(\top \to \top) \to \top} . x(\lambda z^{\top} . x(wz)))\,y : \top$

demo

(STLC)

Induction over Derivations

The Key Idea

The Key Idea

$$\dfrac{}{\{\} \vdash \texttt{2} : \texttt{int}}\ (\text{intLit}) \qquad \dfrac{\dfrac{\dfrac{}{\{\texttt{y} : \texttt{int}\} \vdash \texttt{y} : \texttt{int}}\ (\text{var}) \qquad \dfrac{}{\{\texttt{y} : \texttt{int}\} \vdash \texttt{y} : \texttt{int}}\ (\text{var})}{\{\texttt{y} : \texttt{int}\} \vdash \texttt{y} + \texttt{y} : \texttt{int}}\ (\text{intAdd})}{}$$

$$\dfrac{\{\} \vdash \texttt{2} : \texttt{int} \qquad \{\texttt{y} : \texttt{int}\} \vdash \texttt{y} + \texttt{y} : \texttt{int}}{\{\} \vdash \texttt{let y = 2 in y} + \texttt{y} : \texttt{int}}\ (\text{let})$$

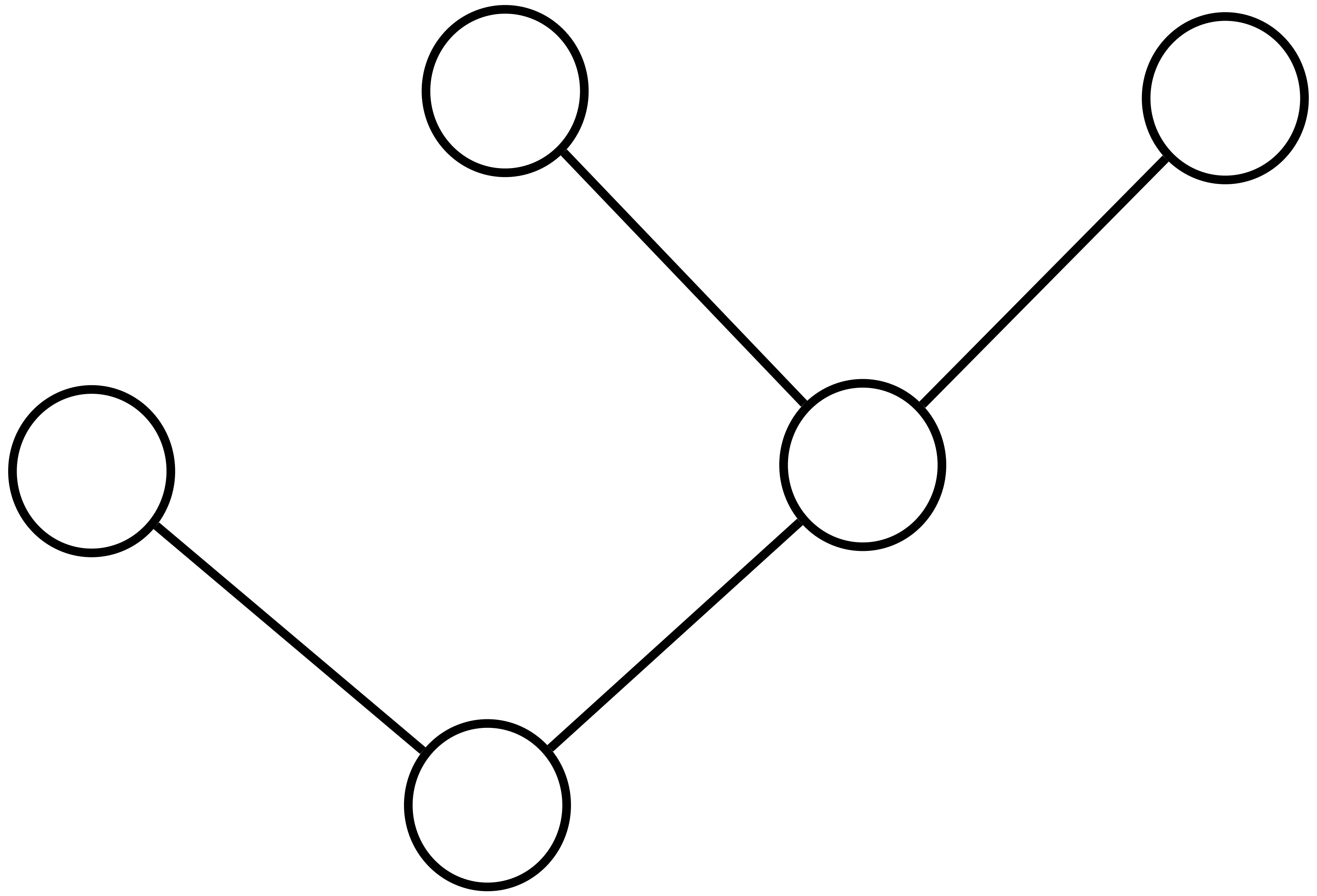
Derivations are *trees*

The Key Idea

$$\dfrac{\dfrac{}{\{\} \vdash \mathtt{2} : \mathtt{int}}\,(\text{intLit}) \qquad \dfrac{\dfrac{}{\{\mathtt{y} : \mathtt{int}\} \vdash \mathtt{y} : \mathtt{int}}\,(\text{var}) \qquad \dfrac{}{\{\mathtt{y} : \mathtt{int}\} \vdash \mathtt{y} : \mathtt{int}}\,(\text{var})}{\{\mathtt{y} : \mathtt{int}\} \vdash \mathtt{y} + \mathtt{y} : \mathtt{int}}\,(\text{intAdd})}{\{\} \vdash \mathtt{let\ y\ =\ 2\ in\ y} + \mathtt{y} : \mathtt{int}}\,(\text{let})$$

Derivations are *trees*

We can prove things about trees using induction



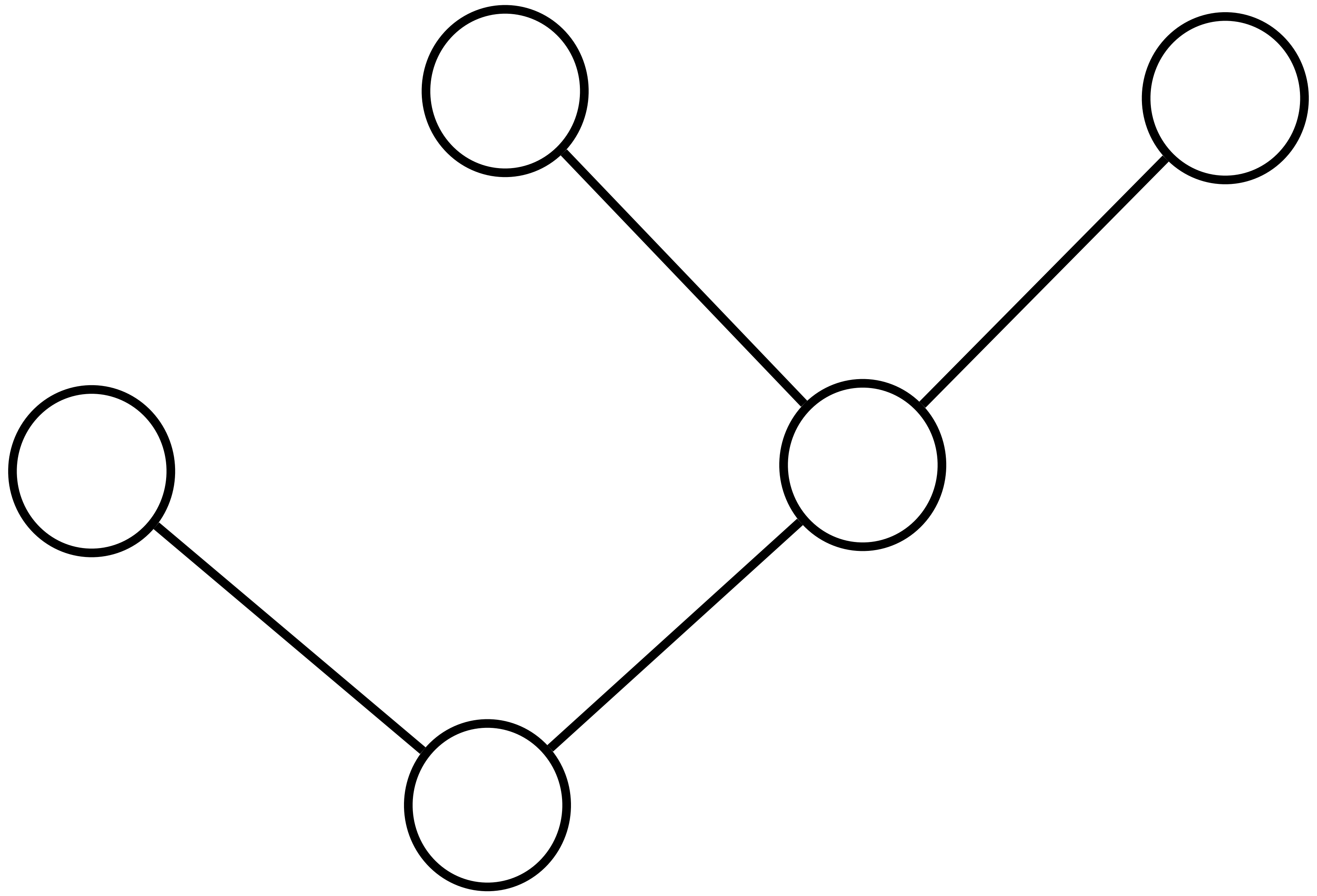
The Key Idea

$$\dfrac{\dfrac{}{\{\} \vdash \texttt{2} : \texttt{int}}\ (\text{intLit}) \qquad \dfrac{\dfrac{}{\{\texttt{y} : \texttt{int}\} \vdash \texttt{y} : \texttt{int}}\ (\text{var}) \qquad \dfrac{}{\{\texttt{y} : \texttt{int}\} \vdash \texttt{y} : \texttt{int}}\ (\text{var})}{\{\texttt{y} : \texttt{int}\} \vdash \texttt{y} + \texttt{y} : \texttt{int}}\ (\text{intAdd})}{\{\} \vdash \texttt{let y = 2 in } \texttt{y} + \texttt{y} : \texttt{int}}\ (\text{let})$$

Derivations are *trees*

We can prove things about trees using induction

We can prove things about *derivable judgments* using induction



The Key Idea

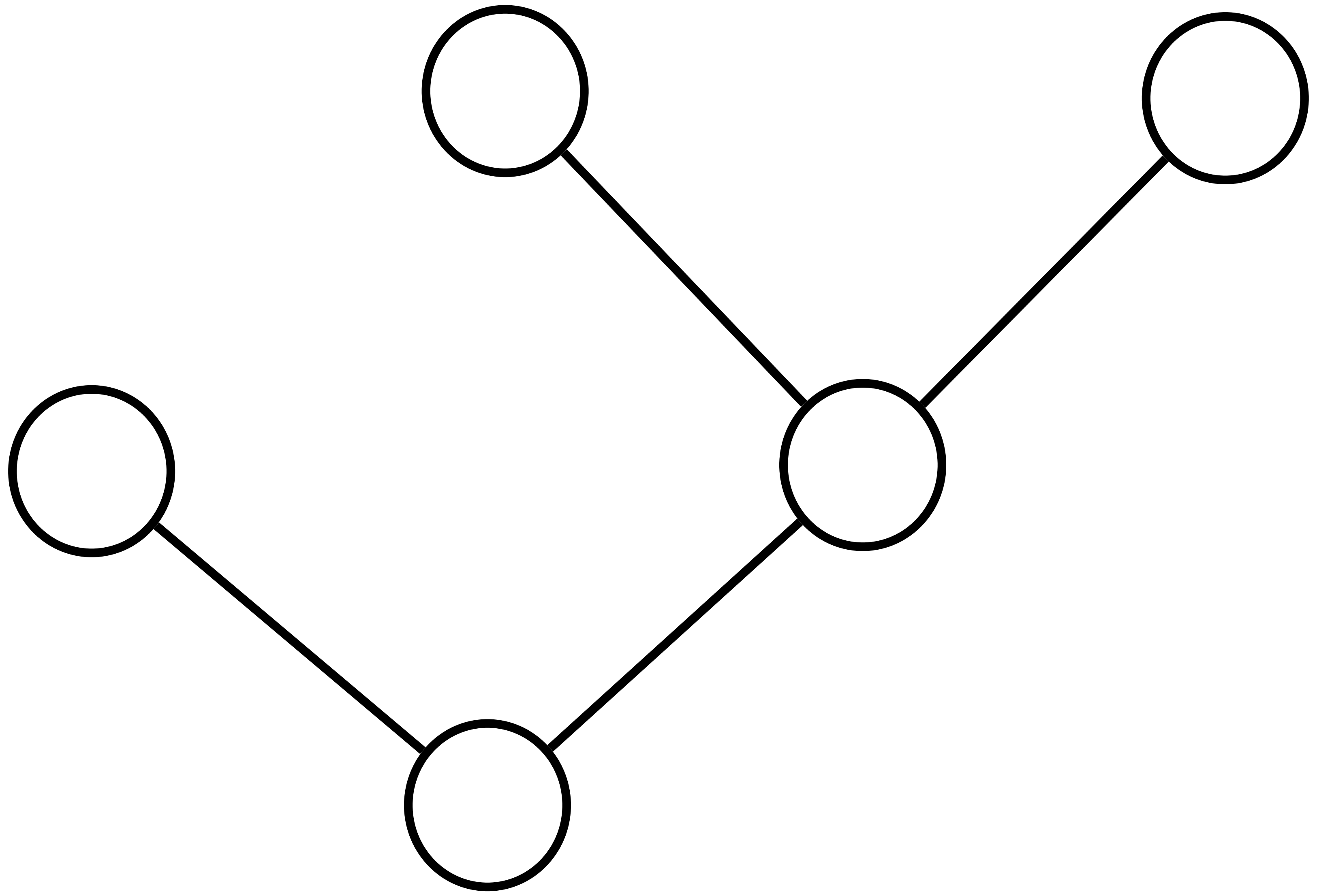
$$\dfrac{\dfrac{}{\{\} \vdash 2 : \texttt{int}}\ (\text{intLit}) \qquad \dfrac{\dfrac{}{\{\texttt{y} : \texttt{int}\} \vdash \texttt{y} : \texttt{int}}\ (\text{var}) \quad \dfrac{}{\{\texttt{y} : \texttt{int}\} \vdash \texttt{y} : \texttt{int}}\ (\text{var})}{\{\texttt{y} : \texttt{int}\} \vdash \texttt{y} + \texttt{y} : \texttt{int}}\ (\text{intAdd})}{\{\} \vdash \texttt{let y = 2 in } \texttt{y} + \texttt{y} : \texttt{int}}\ (\text{let})$$

Derivations are *trees*

We can prove things about trees using induction

We can prove things about *derivable judgments* using induction

Important: Every derivable judgment corresponds to a derivation



Warm-up: Binary Trees

```
type 'a tree =
  | Empty
  | Node of 'a tree * 'a * 'a tree
```

Warm-up: Binary Trees

```
type 'a tree =
  | Empty
  | Node of 'a tree * 'a * 'a tree
```

Let $\text{size}(T)$ denote the number of Nodes and let $\text{height}(T)$ denote the length of the *longest* path from the root to Empty

Warm-up: Binary Trees

```
type 'a tree =
  | Empty
  | Node of 'a tree * 'a * 'a tree
```

Let $\text{size}(T)$ denote the number of Nodes and let $\text{height}(T)$ denote the length of the *longest* path from the root to Empty

Theorem. $\text{size}(T) \leq 2^{\text{height}(T)+1}$ for any tree T

Warm-up: Binary Trees

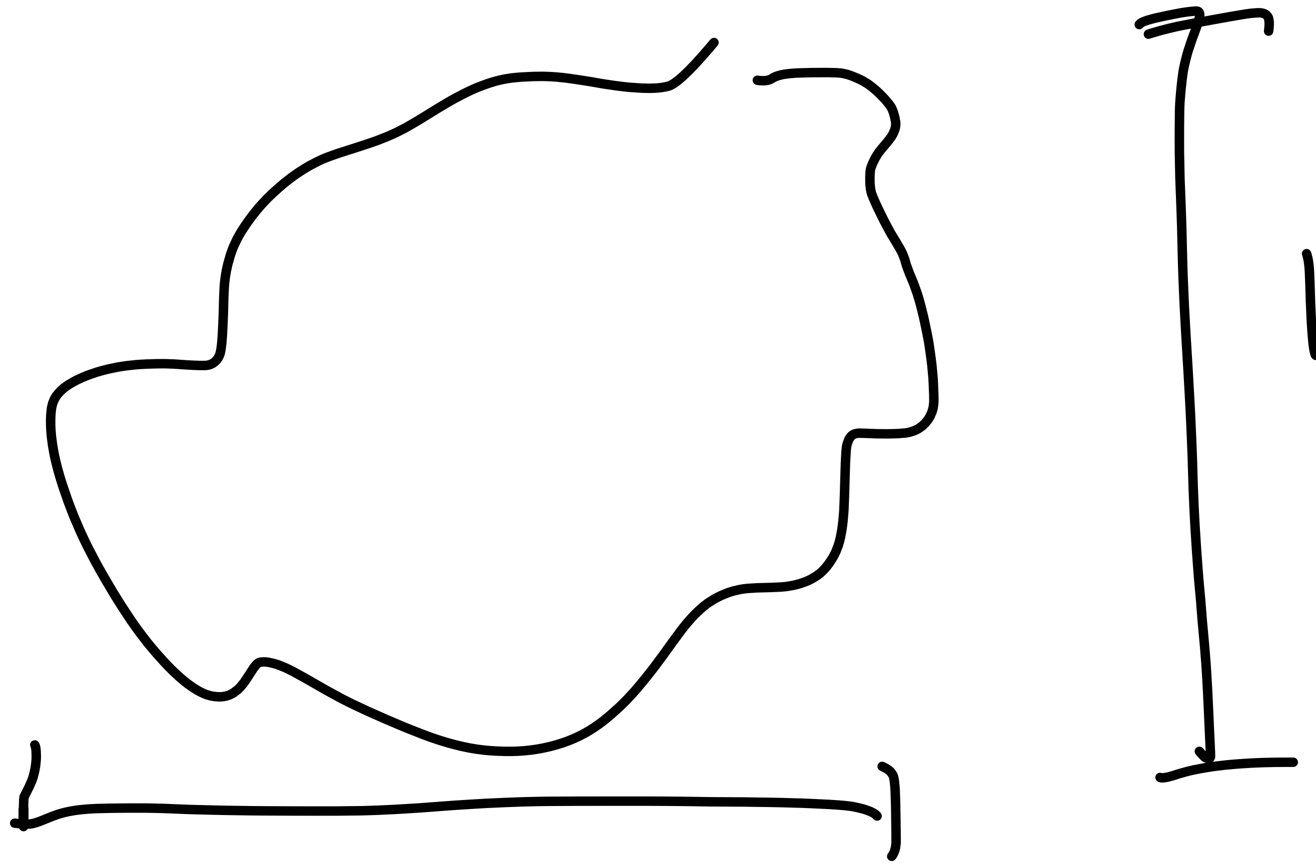
```
type 'a tree =
  | Empty
  | Node of 'a tree * 'a * 'a tree
```

Let $\text{size}(T)$ denote the number of Nodes and let $\text{height}(T)$ denote the length of the *longest* path from the root to Empty

Theorem. $\text{size}(T) \leq 2^{\text{height}(T)}$ for any tree T

Proof. By induction on the structure of trees

Base Case: Empty

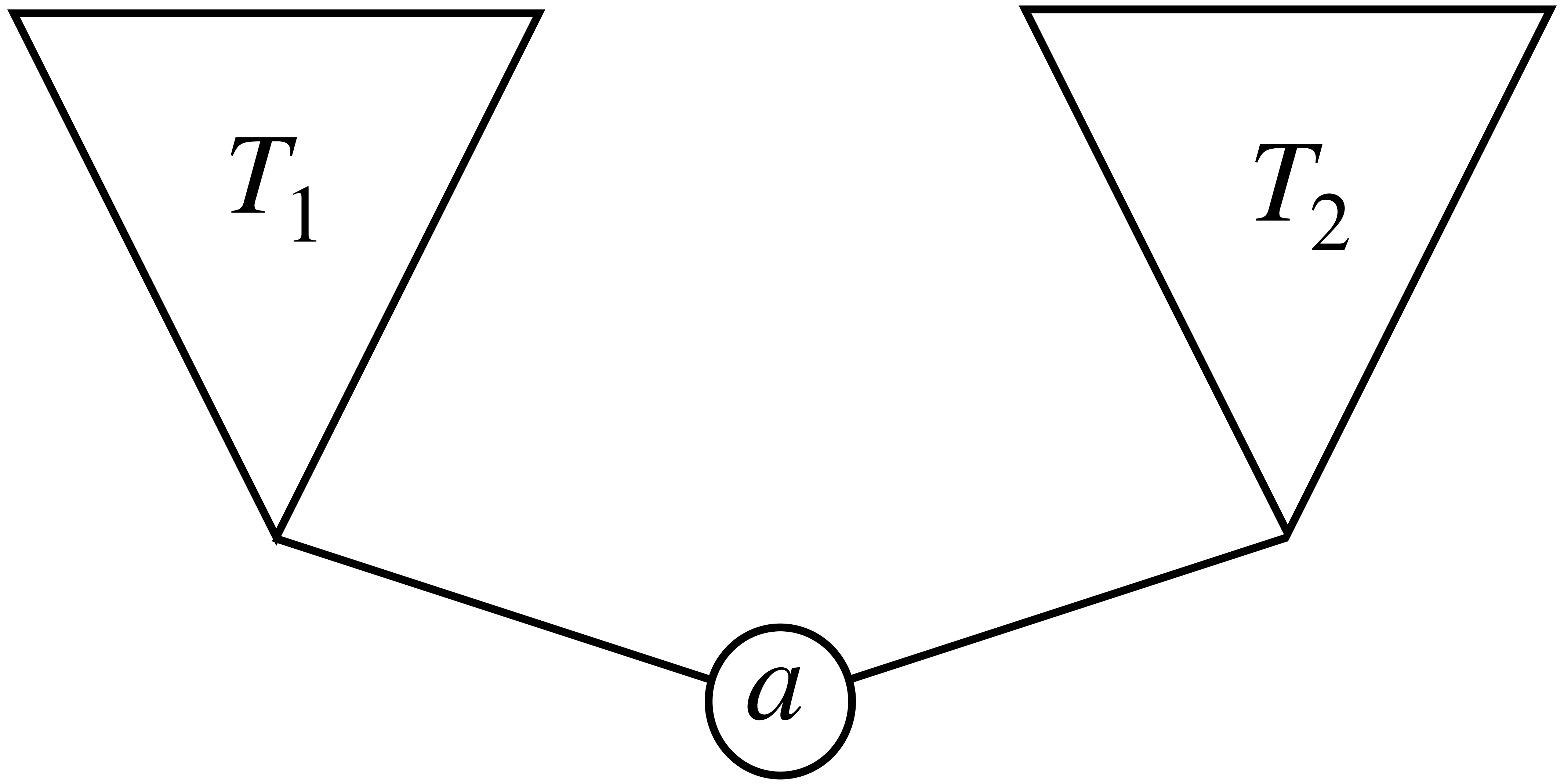


$h = 0$

$n = 0$

$$0 = s(\text{empty}) \leq 2^{h(\text{empty})} = 1$$

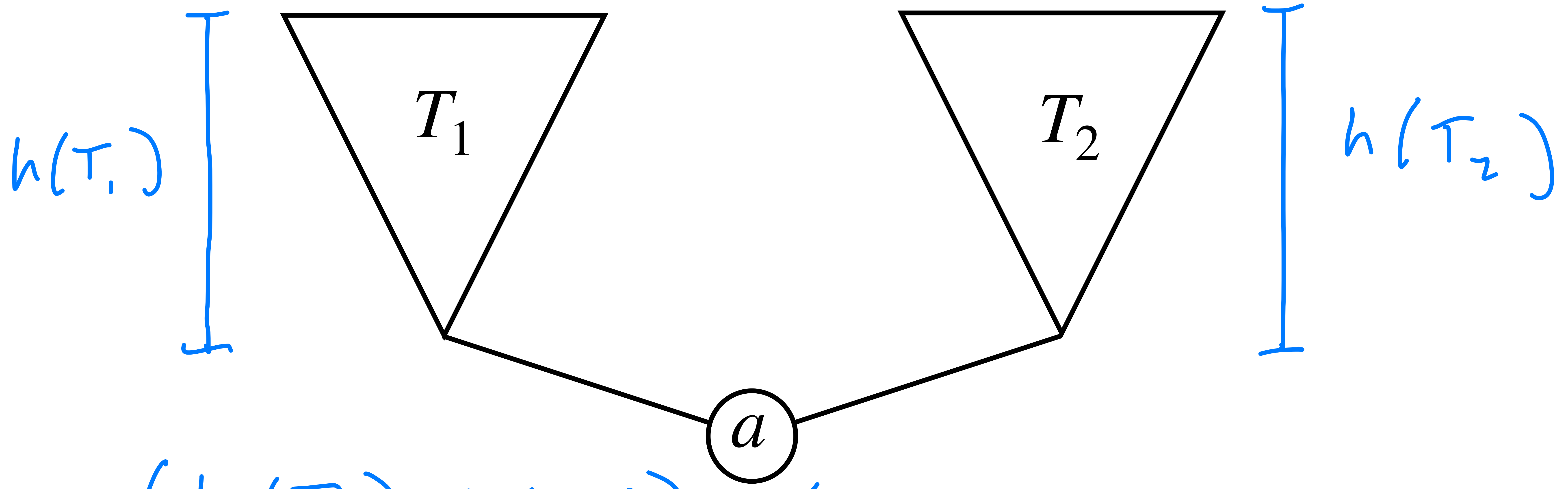
Inductive Hypothesis



If T is `Node` (T_1, a, T_2) then $\text{size}(T_1) \leq 2^{\text{height}(T_1)}$ and $\text{size}(T_2) \leq 2^{\text{height}(T_2)}$

That is, we get to assume that what we want holds of our subtrees

Inductive Step: Nodes



$$h(T) = \max(h(T_1), h(T_2)) + 1$$

$$s(T) = 1 + s(T_1) + s(T_2)$$

$\le 2^{h(T_2)}$ $\quad$ $\le 2^{h(T_1)}$ } by IH

$$1 + \boxed{s(T_1)} + \boxed{s(T_2)} = s(T) \le 2^{h(T)} = 2^{\max(h(T_1), h(T_2)) + 1}$$

Another Warm-up: Well-Scopedness

Another Warm-up: Well-Scopedness

$FV(e) \subseteq Var(\Gamma)$

~~$x : int \vdash y : int$~~

An expression e is **well-scoped** with respect to a context Γ if $x \in FV(e)$ implies x appears in Γ

Another Warm-up: Well-Scopedness

An expression e is **well-scoped** with respect to a context Γ if $x \in FV(e)$ implies x appears in Γ

> Theorem. If e is well-typed in Γ, then e is well-scoped

Another Warm-up: Well-Scopedness

An expression e is **well-scoped** with respect to a context Γ if $x \in FV(e)$ implies x appears in Γ

> Theorem. If e is well-typed in Γ, then e is well-scoped

Proof. By induction on derivations

Base Case: Axioms

$$\frac{}{\Gamma \vdash \bullet : \top}\ \textbf{unit}$$

$FV(\bullet) = \{\ \}$ ✓

$$\frac{(x : \tau) \in \Gamma}{\Gamma \vdash x : \tau}\ \textbf{variable}$$

$FV(x) = \{x\}$ ✓

We need to show that expressions typed using just axioms satisfy well-scopedness

Inductive Hypothesis

$$\frac{\begin{array}{c}\vdots\\ \mathscr{D}_1\\ \boxed{\Gamma_1} \vdash e_1 : \tau_1\end{array} \quad \begin{array}{c}\vdots\\ \mathscr{D}_2\\ \boxed{\Gamma_2} \vdash e_2 : \tau_2\end{array} \quad \ldots \quad \begin{array}{c}\vdots\\ \mathscr{D}_k\\ \boxed{\Gamma_k} \vdash e_k : \tau_k\end{array}}{\Gamma \vdash e : \tau}$$

If $e_1, \ldots, e_k$ are well-scoped (because they are typeable in the each of their contexts)

Inductive Step 1: Application

$$\dfrac{\overset{\vdots}{\mathscr{D}_1} \quad (1)\; \Gamma \vdash e_1 : \tau \to \tau' \qquad \overset{\vdots}{\mathscr{D}_2} \quad \Gamma \vdash e_2 : \tau \;(2)}{\Gamma \vdash e_1 e_2 : \tau'}\ \textbf{application}$$

$$\left.\begin{array}{l} (1)\ FV(e_1) \subseteq Var(\Gamma) \text{ by IH} \\ (2)\ FV(e_2) \subseteq Var(\Gamma) \text{ by IH} \end{array}\right\} \quad \begin{array}{l} FV(e_1 e_2) = FV(e_1) \cup FV(e_2) \\ \text{so} \\ FV(e_1 e_2) \subseteq Var(\Gamma) \end{array}$$

What if the last rule I applied was application?

Inductive Step 2: Abstraction

$$\vdots$$

$$\mathscr{D}$$

(1)

$$\frac{\Gamma, x : \tau \vdash e : \tau'}{\Gamma \vdash \lambda x^{\tau} . e : \tau \rightarrow \tau'} \text{ abstraction}$$

$\checkmark$

$FV(\lambda x^{\tau}.e) \subseteq Var(\Gamma)$

(1) $FV(e) \subseteq Var(\Gamma) \cup \{x\}$ by IH

$\Uparrow$

$FV(\lambda x^{\tau}.e) = FV(e) \setminus \{x\}$, $y \in FV(\lambda x^{\tau}.e) \Rightarrow y \in FV(\Gamma)$ $y \neq x$

What if the last rule I applied was abstraction?

Progress and Preservation

Recall: Type Safety

Recall: Type Safety

> Theorem. If $\cdot \vdash e : \tau$ then there is a value v such that $\langle \varnothing, e \rangle \Downarrow v$ and $\cdot \vdash v : \tau$

Recall: Type Safety

> Theorem. If $\cdot \vdash e : \tau$ then there is a value v such that $\langle \varnothing, e \rangle \Downarrow v$ and $\cdot \vdash v : \tau$

With small-step semantics, we can give a finer-grained analysis:

Recall: Type Safety

> Theorem. If $\cdot \vdash e : \tau$ then there is a value v such that $\langle \varnothing, e \rangle \Downarrow v$ and $\cdot \vdash v : \tau$

With small-step semantics, we can give a finer-grained analysis:

> Theorem. If $\cdot \vdash e : \tau$, then
>
> » *(progress)* either e is a value or there is an e' such that $e \longrightarrow e'$
>
> » *(preservation)* If $\cdot \vdash e : \tau$ and $e \longrightarrow e'$ then $\cdot \vdash e' : \tau$

Recall: Type Safety

> Theorem. If $\cdot \vdash e : \tau$ then there is a value v such that $\langle \varnothing, e \rangle \Downarrow v$ and $\cdot \vdash v : \tau$

With small-step semantics, we can give a finer-grained analysis:

> Theorem. If $\cdot \vdash e : \tau$, then
>
> never get stuck
>
> » *(progress)* either e is a value or there is an e' such that $e \longrightarrow e'$
>
> » *(preservation)* If $\cdot \vdash e : \tau$ and $e \longrightarrow e'$ then $\cdot \vdash e' : \tau$

These results are *fundamental*. They tell us that our PL is well-behaved (it's a "good" PL)

Recall: Type Safety

> Theorem. If $\cdot \vdash e : \tau$ then there is a value v such that $\langle \varnothing, e \rangle \Downarrow v$ and $\cdot \vdash v : \tau$

With small-step semantics, we can give a finer-grained analysis:

goal for today

> Theorem. If $\cdot \vdash e : \tau$, then
>
> » *(progress)* either e is a value or there is an e' such that $e \longrightarrow e'$
>
> » *(preservation)* If $\cdot \vdash e : \tau$ and $e \longrightarrow e'$ then $\cdot \vdash e' : \tau$

These results are *fundamental*. They tell us that our PL is well-behaved (it's a "good" PL)

Disclaimer: We're gonna hand-wave liberally

Recall: STLC

$$e ::= \bullet \mid x \mid \lambda x^{\tau}.e \mid ee$$
$$\tau ::= \top \mid \tau \to \tau$$
$$x ::= \textit{variables}$$

Typing

$$\frac{}{\Gamma \vdash \bullet : \top}\ \texttt{unit} \qquad \frac{\Gamma, x : \tau \vdash e : \tau'}{\Gamma \vdash \lambda x^{\tau}.e : \tau \to \tau'}\ \texttt{abstraction}$$

$$\frac{(x : \tau) \in \Gamma}{\Gamma \vdash x : \tau}\ \texttt{variable} \qquad \frac{\Gamma \vdash e_1 : \tau \to \tau' \qquad \Gamma \vdash e_2 : \tau}{\Gamma \vdash e_1 e_2 : \tau'}\ \texttt{application}$$

Semantics

$$\frac{e_1 \longrightarrow e_1'}{e_1 e_2 \longrightarrow e_1' e_2}\ \texttt{leftEval} \qquad \frac{}{(\lambda x.e)e' \longrightarrow [e'/x]e}\ \texttt{beta}$$

Progress (STLC)

> **Theorem.** If e is well-typed ($\boxed{\cdot \vdash e : \tau}$ for some type τ), then e is a value, or there is an expression e' such that $e \longrightarrow e'$

Proof. By induction over derivations

Base Case: Axioms

$$\frac{}{\cdot \vdash \bullet : \top}\ \textbf{unit}$$

value ✓

$$\cancel{\frac{(x : \tau) \in \varnothing}{\cdot \vdash x : \tau}}\ \textbf{variable}$$

empty

impossible : well-scoped

We need to show that expressions typed using just axioms yield non-stuck terms

Inductive Step 1: Application

(1)

$$\frac{\cdot \vdash e_1 : \tau \to \tau' \qquad \cdot \vdash e_2 : \tau}{\cdot \vdash e_1 e_2 : \tau'} \ \textbf{application}$$

(1) e_1 is either a value (a) $(\lambda x^{\tau}. e)$ or (b) $e \to e'$ for some e'

(a) $$\frac{}{(\lambda x^{\tau} e) e_2 \to [e_2 / x] e} \ (\text{beta})$$

(b) $e_1 \to e_1'$ implies $e_1 e_2 \to e_1' e_2$ (appleft)

*What do we know given that e_1 is either a **value** or **reducible**?*

Inductive Step 2: Abstraction

$$\dfrac{\{x:\tau\} \vdash e : \tau'}{\cdot \vdash \underbrace{\lambda x^{\tau}.e}_{\text{value}} : \tau \rightarrow \tau'}\ \textbf{abstraction}$$

Our expression already a value if the last rule we applied was abstraction!

Preservation (STLC)

> Theorem. If e has type τ in Γ (i.e., $\Gamma \vdash e : \tau$ is derivable) and $e \longrightarrow e'$ then so is e' (i.e., $\Gamma \vdash e' : \tau$ is derivable)

Proof. By induction over derivations

This one is much tricker...

Base Case: Axioms

$$\frac{}{\Gamma \vdash \bullet : \top}\,\texttt{unit}$$

$$\frac{(x : \tau) \in \Gamma}{\Gamma \vdash x : \tau}\,\texttt{variable}$$

Expressions typed using just axioms cannot be reduced (nothing to do here)

Inductive Step 1: Abstraction

$$\frac{\Gamma, x : \tau \vdash e : \tau'}{\Gamma \vdash \lambda x^{\tau} . e : \tau \to \tau'} \;\texttt{abstraction}$$

Expressions derived using abstraction as the last rule is already a value (nothing to do here)

Inductive Step 2: Application

$$\frac{\Gamma \vdash e_1 : \tau \to \tau' \qquad \Gamma \vdash e_2 : \tau}{\Gamma \vdash e_1 e_2 : \tau'}\ \texttt{application}$$

This is where the work comes in...

<u>The trick:</u> We do induction (inside our current induction) on the structure of *semantic* derivations!

What possible ways can $e_1 e_2$ *be reduced?*

Inductive Step 2.1: leftEval

$$\frac{e_1 \longrightarrow e_1'}{e_1 e_2 \longrightarrow e_1' e_2}\ \texttt{leftEval} \qquad \frac{\Gamma \vdash e_1 : \tau \rightarrow \tau' \qquad \Gamma \vdash e_2 : \tau}{\Gamma \vdash e_1 e_2 : \tau'}\ \texttt{application}$$

*What if our last rule was an application **and** $e_1 e_2$ is reducible by leftEval?*

Inductive Step 2.1: leftEval

$$\frac{}{(\lambda x\,.\,e)e_2 \longrightarrow [e_2/x]e}\ \textbf{beta}$$

$$\frac{\Gamma \vdash (\lambda x\,.\,e) : \tau \to \tau' \qquad \Gamma \vdash e_2 : \tau}{\Gamma \vdash (\lambda x\,.\,e)e_2 : \tau'}\ \textbf{application}$$

What if our last rule was an application ***and*** $e_1 e_2$ *is reducible by beta?*

Substitution Lemma

Lemma. If $\Gamma \vdash e_2 : \tau_2$ and $\Gamma, x : \tau_2 \vdash e : \tau$ then

$$\Gamma \vdash [e_2/x]e : \tau$$

That is, if e is well-typed in a context with (x, τ) then we can substitute x with anything of type τ and it's still the same type

(we can prove this by, you guessed it, induction on derivations)

The Point

```
let rec eval env e =
  match e with
  | Var x -> Env.find x env
  ...
```

fails if not found

Progress and preservation tell us that **terms never get stuck during evaluation**

*This is **HUGE**. I can't emphasize this enough*

Our type system ensures we only evaluate programs that make sense!

Summary

» **Progress** and **preservation** are fundamental features of good programming languages

» We can prove things about well-typed expressions by performing **induction** over derivations