Normalization of $\beta$

Weak normalization (WN)
A term $M$ is WN if there exists a reduction
$M \rightarrow_\beta M_1 \rightarrow_\beta \cdots \rightarrow_\beta M_n$
where $M_n$ is in normal form.
Normalization of $\beta$

**Weak normalization (WN)**
A term $M$ is WN if there exists a reduction

$$M \rightarrow^{\beta} M_1 \rightarrow^{\beta} \cdots \rightarrow^{\beta} M_n$$

where $M_n$ is in normal form

**Strong normalization (SN)**
A term $M$ is SN if there exist no infinite reductions starting from $M$

i.e. all $\beta$-reductions from $M$ lead to a normal form

i.e. there exists a bound on the length of $\beta$-reductions from $M$

(because $\rightarrow^{\beta}$ is finitely branching)
A naive proof attempt

**Theorem:** \( \Gamma \vdash M : \sigma \) implies \( M \in SN \).

Proof. By structural induction on \( M \).

- **Case** \( M = x \): A variable cannot be reduced.

- **Case** \( M = \lambda x . N \): From the hypothesis, \( \Gamma \vdash \sigma \rightarrow \tau \). Hence \( \Gamma \vdash \sigma \). By IH, \( N \in SN \). Therefore \( M \in SN \).

- **Case** \( M = M_1 M_2 \): We have \( \Gamma \vdash M_1 M_2 : \tau \). Hence \( \Gamma \vdash M_1 : \sigma \rightarrow \tau \) and \( \Gamma \vdash M_2 : \sigma \). By the IHs, we have \( M_1 \in SN \) and \( M_2 \in SN \), but this doesn't tell us anything about \( M_1 M_2 \)—e.g. maybe \( M_1 \) reduces to \( \lambda x . N_1 \), yielding a new redex. Indeed, try \( M_1 = M_2 = (\lambda x . x x) \).
A naive proof attempt

**Theorem:** \( \Gamma \vdash M : \sigma \) implies \( M \in SN \).

**Proof.** By structural induction on \( M \).
A naive proof attempt

**Theorem:** $\Gamma \vdash M : \sigma$ implies $M \in SN$.

**Proof.** By structural induction on $M$.

**Case $M = x$:** A variable cannot be reduced.
A naive proof attempt

**Theorem:** $\Gamma \vdash M : \sigma$ implies $M \in SN$.

**Proof.** By structural induction on $M$.

**Case $M = x$:** A variable cannot be reduced.

**Case $M = \lambda x. N$:** From the hypothesis, $\Gamma \vdash \sigma \rightarrow \tau$. Hence $\Gamma \vdash \sigma$. By IH, $N \in SN$. Therefore $M \in SN$. 

Indeed, try $M_1 = M_2 = (\lambda x. x x)$.
A naive proof attempt

**Theorem:** $\Gamma \vdash M : \sigma$ implies $M \in SN$.

**Proof.** By structural induction on $M$.

**Case $M = x$:** A variable cannot be reduced.

**Case $M = \lambda x. N$:** From the hypothesis, $\Gamma \vdash \sigma \rightarrow \tau$. Hence $\Gamma \vdash \sigma$. By IH, $N \in SN$. Therefore $M \in SN$.

**Case $M = M_1M_2$:** We have $\Gamma \vdash M_1M_2 : \tau$. Hence $\Gamma \vdash M_1 : \sigma \rightarrow \tau$ and $\Gamma \vdash M_2 : \sigma$. By the IHs, we have $M_1 \in SN$ and $M_2 \in SN$, but this doesn’t tell us anything about $M_1M_2$—e.g. maybe $M_1$ reduces to $\lambda x. N_1$, yielding a new redex.  

□
A naive proof attempt

**Theorem:** \( \Gamma \vdash M : \sigma \) implies \( M \in SN \).

**Proof.** By structural induction on \( M \).

**Case** \( M = x \): A variable cannot be reduced.

**Case** \( M = \lambda x. N \): From the hypothesis, \( \Gamma \vdash \sigma \rightarrow \tau \). Hence \( \Gamma \vdash \sigma \). By IH, \( N \in SN \). Therefore \( M \in SN \).

**Case** \( M = M_1 M_2 \): We have \( \Gamma \vdash M_1 M_2 : \tau \). Hence \( \Gamma \vdash M_1 : \sigma \rightarrow \tau \) and \( \Gamma \vdash M_2 : \sigma \). By the IHs, we have \( M_1 \in SN \) and \( M_2 \in SN \), but this doesn’t tell us anything about \( M_1 M_2 \)---e.g. maybe \( M_1 \) reduces to \( \lambda x. N_1 \), yielding a new redex.

Indeed, try \( M_1 = M_2 = (\lambda x. x x) \).
Difficulties

- Terms may get **larger** under reduction:
  \[(\lambda f. \, g \, f \, f) \text{BIG} \rightarrow_\beta g \, \text{BIG BIG}\]

- Redexes may get **multiplied** under reduction:
  \[(\lambda f. \, g \, f \, f) \, ((\lambda x. \, M)Q) \rightarrow_\beta g \, ((\lambda x. \, M)Q) \, ((\lambda x. \, M)Q)\]

- New redexes can be **created** under reduction:
  \[(\lambda f. \, f \, x) \, (\lambda y. \, N) \rightarrow_\beta (\lambda y. \, N) \, x\]
Difficulties

- Terms may get larger under reduction:
  \[(\lambda f. \ g \ f \ f) \text{BIG} \rightarrow_\beta g \ \text{BIG \ BIG}\]

- Redexes may get multiplied under reduction:
  \[(\lambda f. \ g \ f \ f)((\lambda x. \ M)Q) \rightarrow_\beta g ((\lambda x. \ M)Q)((\lambda x. \ M)Q)\]

- New redexes can be created under reduction:
  \[(\lambda f. \ f \ x)(\lambda y. \ N) \rightarrow_\beta (\lambda y. \ N) \ x\]

Match plan:
1. Prove WN (following Turing, Gandy)
2. Prove SN (following Tait)
Redex creation

In the (untyped) $\lambda$-calculus, there are three ways to create “new” $\beta$-redexes.

- **Substitution:**
  \[(\lambda x. \ldots (x \ P) \ldots) (\lambda y. \ Q) \to_\beta \ldots (\lambda y. \ Q) \ P \ldots\]

- **Multiplication:**
  \[(\lambda x. \ldots x \ldots x \ldots) ((\lambda y. \ Q) \ R) \to_\beta \ldots((\lambda y. \ Q) \ R) \ldots((\lambda y. \ Q) \ R) \ldots\]

- **Identity:**
  \[(\lambda x. \ x) (\lambda y. \ Q) \ R \to_\beta (\lambda y. \ Q) \ R\]
Height (cf. order)

\[ h(a) = 0 \]
\[ h(\sigma_1 \rightarrow \cdots \rightarrow \sigma_n \rightarrow a) = \max\{h(\sigma_1), \ldots, h(\sigma_n)\} + 1 \]

i.e. \( h(\sigma \rightarrow \tau) = \max\{h(\sigma) + 1, h(\tau)\} \)

The height of a redex \((\lambda x. P) Q\) is the height of the type of \(\lambda x. P\)
We define a **measure** $m$ as follows:

$$m(N) = (h(N), \#N)$$

where

$h(N) =$ the maximum height of a redex in $N$

$\#N =$ the number of redexes of height $h(N)$ in $N$
We define a measure $m$ as follows:

$$m(N) = (h(N), \#N)$$

where

$h(N) =$ the maximum height of a redex in $N$

$\#N =$ the number of redexes of height $h(N)$ in $N$

The measures are ordered lexicographically:

$$(h', n') > (h, n) \iff h' > h \text{ or } h' = h \text{ and } n' > n$$
Measure

We define a measure $m$ as follows:

$$m(N) = (h(N), \#N)$$

where

$h(N)$ = the maximum height of a redex in $N$

$\#N$ = the number of redexes of height $h(N)$ in $N$

The measures are ordered lexicographically:

$$(h', n') > (h, n) \text{ iff } h' > h \text{ or } h' = h \text{ and } n' > n$$

Fact: $>$ on measures is well founded
Weak normalization

**Theorem:** If \( P \) is typable in \( \lambda \rightarrow \), then there is a terminating reduction starting from \( P \).
Theorem: If $P$ is typable in $\lambda\rightarrow$, then there is a terminating reduction starting from $P$.

Proof.
Weak normalization

**Theorem:** If $P$ is typable in $\lambda \rightarrow$, then there is a terminating reduction starting from $P$.

**Proof.**

Pick a redex of height $h(P)$ inside $P$ that does not contain any other redex of height $h(P)$. This is always possible.

Contract this redex, yielding $Q$.

This *does not create a new redex of height* $h(P)$. Consider the three ways in which redexes can be created.
Strong computability (SC)

\[ M : a \in SC \iff M \in SN \]
\[ M : \sigma \rightarrow \tau \in SC \iff \text{for all } N : \sigma \in SC, \ M \ N \in SC \]
Lemma 1

(a) $\times N_1 \ldots N_k : \sigma \in SC$ if $N_1, \ldots, N_k \in SN$.

(b) $M \in SC$ implies $M \in SN$ for $M : \sigma$. 
Lemma 2

Let $N \in SC$. If $M[x := N] \in SC$, then $(\lambda x. M) N \in SC$. 
Lemma 3 and Corollary

Let $\Gamma = x_1 : \tau_1, \ldots, x_n : \tau_n$.
Assume $\Gamma \vdash M : \sigma, N_1 : \tau_1, \ldots, N_n : \tau_n$ and $N_1, \ldots, N_n \in SC$.
Then $M[(x_1, \ldots, x_n) := (N_1, \ldots, N_n)] \in SC$.

Corollary: $\lambda \rightarrow$ is SN.
Lemma 3 and Corollary

Let $\Gamma = x_1 : \tau_1, \ldots, x_n : \tau_n$.
Assume $\Gamma \vdash M : \sigma, N_1 : \tau_1, \ldots, N_n : \tau_n$ and $N_1, \ldots, N_n \in SC$.
Then $M[(x_1, \ldots, x_n) := (N_1, \ldots, N_n)] \in SC$.

Corollary: $\lambda \rightarrow$ is SN.