Lesson 5: CFG = PDA

Theme: The equivalence between context-free grammars (CFG) and push-down automata (PDA).

We are going to show that CFG and PDA define precisely the same class of languages. More precisely, we are going to show the following.

- For every CFG $G$, there is a PDA $A$ such that $L(A) = L(G)$.
- For every PDA $A$, there is a CFG $G$ such that $L(A) = L(G)$.

1 From CFG to PDA

Allowing the PDA to push a string of symbols. First, we note that we can modify the definition of CFG to one that allows it to push a string of symbols to its stack. That is, the transitions can be of the form:

$$(p, x, \text{pop}(y)) \rightarrow (q, \text{push}(z)),$$

where $z \in \Gamma^*$.

Allowing such transition does not change the capability of a CFG. We can add “new” states $t_1, \ldots, t_m$, where $m$ is the length of $z$ and $z = a_1 \ldots a_m$. Each $t_i$ is used to push the symbol $a_i$ into the stack. More formally, the transition $(p, x, \text{pop}(y)) \rightarrow (q, \text{push}(z))$ can be replaced with the following transitions:

- $(p, x, \text{pop}(y)) \rightarrow (t_1, \text{push}(a_1))$
- $(t_1, \epsilon, \text{pop}(\epsilon)) \rightarrow (t_2, \text{push}(a_2))$
- $(t_2, \epsilon, \text{pop}(\epsilon)) \rightarrow (t_3, \text{push}(a_3))$
  
  $\vdots$
  
  $(t_{i-1}, \epsilon, \text{pop}(\epsilon)) \rightarrow (t_i, \text{push}(a_i))$
  
  $\vdots$
  
- $(t_{m-1}, \epsilon, \text{pop}(\epsilon)) \rightarrow (t_m, \text{push}(a_m))$
- $(t_m, \epsilon, \text{pop}(\epsilon)) \rightarrow (q, \text{push}(\epsilon))$

Left-most substitution property. Let $G = \langle \Sigma, V, R, S \rangle$ be a CFG. Let $z_1 \Rightarrow z_2 \Rightarrow \cdots \Rightarrow z_n$ be a derivation. We say that the derivation has the left-most substitution property, if for every $i \in \{1, \ldots, n-1\}$, $z_i \Rightarrow z_{i+1}$ is obtained by applying a rule $A \rightarrow w$, where $A$ is the left-most variable in $z_i$. Intuitively, it means that we only substitute the left-most variable in our derivation.

For example, suppose we have grammars with the following rules: $A \rightarrow aABa$, $A \rightarrow SS$, $B \rightarrow aab$, $B \rightarrow SA$. The derivation $aABAaa \Rightarrow aSSBAaa$ has the left-most substitution property, because we substitute the left-most variable which is $A$. On the other hand, $aABAaa \Rightarrow aASAAaa$ does not, because we substitute variable $B$, which is not the left-most variable, and $aABAaa \Rightarrow aABSSaa$ does not either, because we substitute the second $A$, which is not the left-most.

Remark 5.1 Without loss of generality, we only need to consider derivations that has left-most substitution property, by simply substituting the left-most variable first.
Constructing a PDA from a CFG. Let $G = \langle \Sigma, V, R, S \rangle$ be a CFG. Consider the following PDA $A = \langle \Sigma, \Gamma, Q, q_0, F, \delta \rangle$ where each component is as follows.

- $\Gamma = \Sigma \cup V \cup \{\perp\}$
- $Q = \{p, q, r\}$
- $p$ is the initial state.
- $F = \{r\}$
- $\delta$ consists of the following transitions. (Here $w^r$ denotes the reverse of $w$.)
  - $(p, \epsilon, \text{pop}(\epsilon)) \rightarrow (q, \text{push}(\perp S))$
  - $(q, a, \text{pop}(a)) \rightarrow (q, \text{push}(\epsilon))$, for every $a \in \Sigma$.
  - $(q, \epsilon, \text{pop}(A)) \rightarrow (q, \text{push}(s^r))$, for every rule $A \rightarrow s \in R$.
  - $(q, \epsilon, \text{pop}(\perp)) \rightarrow (r, \text{push}(\epsilon))$.

Notice that in the first and third transitions, $A$ pushes a string of symbols to its stack.

We will show that $L(G) = L(A)$. First, we show the following lemma.

**Lemma 5.2** For every derivation:

$$u \Rightarrow \tilde{u}$$

there is a run of $A$ on $w$:

$$(q, \perp v) \vdash_w^* (q, \perp \tilde{v})$$

where

$$v^r = u \quad \text{and} \quad w\tilde{v}^r = \tilde{u} \quad \quad (†)$$

**Proof.** Suppose $u \Rightarrow^* \tilde{u}$. Let $m$ be the length of the derivation, which is denoted as follows.

$$u_0 \Rightarrow u_1 \Rightarrow \cdots \Rightarrow u_{m-1} \Rightarrow u_m$$

where $u_0 = u$ and $u_m = \tilde{u}$. We will prove the lemma by induction on $m$.

The base case is $m = 0$. In this case the derivation is a trivial derivation:

$$u_0 \Rightarrow^* u_0$$

The run:

$$(q, \perp v) \vdash_\epsilon^* (q, \perp v), \quad \text{where} \quad v^r = u_0$$

satisfies property $(†)$.

For the induction hypothesis, we assume that the lemma holds when the length of the derivation is $m - 1$. We will prove the derivation of length $m$ case for the induction step.

Consider a derivation as in $(†)$. We may assume that it has the left-most substitution property. Suppose the derivation $u_0 \Rightarrow u_1$ is obtained by applying the rule $A \rightarrow s$, where variable $A$ is the left-most variable in $u_0$. So, we can denote $u_0$ by $xAy$, where $x \in \Sigma^*$, and hence, $u_1 = xsy$. Since $x$ contains only terminals, using transitions of the form $(q, a, \text{pop}(a)) \rightarrow (q, \text{push}(\epsilon))$, where $a \in \Sigma$, there is a run:

$$(q, \perp y^r Ax^r) \vdash_x^* (q, \perp y^r A). \quad (2)$$
By our construction of $A$, there is a transition $(q, \epsilon, \text{pop}(A)) \rightarrow (q, \text{push}(s'))$. So, we have:

$$ (q, \bot y^r A) \vdash_\epsilon (q, \bot y^r s'). $$

(3)

Moreover, $x$ will not change during the derivation $u_0 \Rightarrow u_1 \Rightarrow \cdots \Rightarrow u_m$ (because $x$ contains only terminals). Thus, the string $u_i$ has prefix $x$, for every $i \in \{1, \ldots, m\}$. That is, $u_i = xu'_i$, for some $u'_i \in (\Sigma \cup V)^*$. So, the derivation is of the form:

$$xAy \Rightarrow xu'_1 \Rightarrow \cdots \Rightarrow xu'_m.$$ 

Ignoring the first part of the derivation, we have:

$$ xu'_1 \Rightarrow \cdots \Rightarrow xu'_m, $$

which means we also have derivation (since $x$ contains only terminals):

$$u'_1 \Rightarrow \cdots \Rightarrow u'_m,$$

which has length $m - 1$. By the induction hypothesis, there is a run:

$$ (q, \bot z_1) \vdash^*_t (q, \bot z_2), $$

where $z_1^r = u'_1$ and $tz_2^r = u'_{m+1}$.

(4)

Now, recall that $u_1 = xu'_1 = xsy$, and hence, $z_1 = y^rs'$. Combining the runs (2), (3) and (4), we have the following run:

$$ (q, \bot y^r Ax^r) \vdash^*_x (q, \bot y^r A) \vdash_\epsilon (q, \bot y^r s') \vdash^*_t (q, \bot z_2) $$

which can be abbreviated as:

$$ (q, \bot y^r Ax^r) \vdash^*_xt (q, \bot z_2). $$

Since $tz_2^r = u'_m$, we have $xtz_2^r = xu'_m = u_m$. Since $u_0 = xAy$, this is the desired property (†). This completes the proof of Lemma 5.2.

The next lemma is the converse direction of Lemma 5.2.

**Lemma 5.3** For every run of $A$:

$$ (q, \bot v) \vdash^*_w (q, \bot \bar{v}), $$

there is a derivation:

$$ u \Rightarrow^* \tilde{u} $$

such that:

$$ v^r = u \quad \text{and} \quad w\bar{v}^r = \tilde{u} \quad (\ast) $$

**Proof.** Suppose we have a run $(q, \bot v) \vdash^*_w (q, \bot \bar{v})$. Let $n$ be the length of the run, which is denoted as follows.

$$ (q, \bot v_0) \vdash_{b_1} (q, \bot v_1) \vdash_{b_2} \cdots \vdash_{b_n} (q, \bot v_n), $$

(5)

where $v_0 = v$, $v_n = \bar{v}$ and $b_1 \cdots b_n = w$. We will prove the lemma by induction on $n$.

The base case is $n = 0$. In this case, the run is a trivial run:

$$ (q, \bot v_0) \vdash_\epsilon (q, \bot v_0) $$
Thus, the trivial derivation:

\[ u \Rightarrow^* u \]  

where \( u = v_0^r \)

satisfies the desired property (\( \star \)).

For the induction hypothesis, we assume that the lemma holds when the length of the run is \( n - 1 \). We will prove the case when the length of the run is \( n \) for the induction step.

We consider a run of length \( n \) as in (5). We consider the step \((q, \perp) \vdash b_1 (q, \perp)\). There are two cases:

- \( b_1 \neq \epsilon \), i.e., \( b_1 \) is a symbol from \( \Sigma \).

  By the construction of the PDA \( A \), it has to use the transition \((q, b_1, \text{pop}(b_1)) \rightarrow (q, \text{push}(\epsilon))\), which means that the top of the stack \( v_0 \) is \( b_1 \). Therefore, \( v_0 = v_1 b_1 \).

  Now consider the run:

  \[(q, \perp) \vdash b_2 \cdots \vdash b_n (q, \perp)\].

  This run is of length \( n - 1 \). By the induction hypothesis, we have a derivation:

  \[ u_1 \Rightarrow^* u_2 \]  

  where \( u_1 = v_1^r \) and \( u_2 = b_2 \cdots b_n v_n^r \).

  Adding \( b_1 \) in front of \( u_1 \) and \( u_2 \), we have:

  \[ b_1 u_1 \Rightarrow^* b_1 u_2 \]

  Now, \( b_1 u_1 = b_1 v_1^r = v_0^r \) and \( b_1 u_2 = b_1 b_2 \cdots b_n v_n^r \), which is the desired property (\( \star \)).

- \( b_1 = \epsilon \).

  By the construction of the PDA \( A \), it has to use the transition \((q, \epsilon, \text{pop}(A)) \rightarrow (q, \text{push}(s^r))\), for some rule \( A \rightarrow s \). This means that the top of the stack \( v_0 \) is \( A \). We denote by \( v_0 = yA \) and \( v_1 = ys^r \), for some \( y \).

  Consider the run:

  \[(q, \perp) \vdash b_2 \cdots \vdash b_n (q, \perp)\].

  This run has length \( n - 1 \). By the induction hypothesis, there is a derivation:

  \[ u_1 \Rightarrow^* u_2 \]  

  where \( u_1 = v_1^r \) and \( u_2 = b_2 \cdots b_n v_n^r \).

  Since \( v_1 = ys^r \), and hence, \( u_1 = sy^r \), we have a derivation:

  \[ Ay^r \Rightarrow u_1 \Rightarrow^* u_2. \]

  By our notation, \( v_0 = yA \) and \( u_2 = b_2 \cdots b_n v_n^r = b_1 b_2 \cdots b_n v_n^r \), which is the desired property (\( \star \)).

This completes our proof of Lemma 5.3. ■

Using the two lemmas above, we can show that \( L(A) = L(G) \) as stated below.

**Theorem 5.4** \( L(G) = L(A) \).
Proof. We first prove $L(G) \subseteq L(A)$. Let $S \Rightarrow^* w$, where $w \in \Sigma^*$. By Lemma 5.2 there is a run:

\[(q, \bot S) \vdash^*_w (q, \bot y), \quad \text{where } xy^r = w.\]

Since $w$ contains only symbols from $\Sigma$, so does $y$. If $y \neq \epsilon$, using the transitions of the form $(q, a, \text{pop}(a)) \rightarrow (q, \text{push}(\epsilon))$, we can extend the run until $y = \epsilon$, which means that $x = w$.

By the construction of $A$, we have the following accepting run on $w$:

\[(p, \epsilon) \vdash^*(q, \bot S) \vdash^*_w (q, \bot) \vdash^* (r, \epsilon).\]

Thus, $w \in L(A)$.

Now we prove $L(G) \supseteq L(A)$. Let $w \in L(A)$. So, there is an accepting run of $A$ on $w$, which must start from the initial state $p$ and end with the accepting state $r$. Thus, it has to start by using the transition $(p, \epsilon, \text{pop}(\bot)) \rightarrow (q, \text{push}(\bot S))$ and end by using the transition $(q, \epsilon, \text{pop}(\bot)) \rightarrow (r, \text{push}(\epsilon))$. This means the run is of the form:

\[(p, \epsilon) \vdash^*(q, \bot S) \vdash^*_w (q, \bot) \vdash^* (r, \epsilon).\]

By Lemma 5.3 there is a derivation:

\[u_1 \Rightarrow^* u_2, \quad \text{where } u_1 = S \text{ and } u_2 = w.\]

Thus, $S \Rightarrow^* w$ and $w \in L(G)$. This completes the proof of Theorem 5.4.

\[\Box\]

2 From PDA to CFG

Let $A = (\Sigma, \Gamma, Q, q_0, F, \delta)$ be a PDA. Without loss of generality, we can assume the following.

- It has only one final state, say $q_f$. That is, $F = \{q_f\}$.
- The stack is empty before accepting an input word. More precisely, on every word $w$, if $A$ accepts $w$, there is an accepting run of $A$ on $w$ from the initial configuration $(q_0, \epsilon)$ to a final configuration $(q_f, \epsilon)$ where the content of the stack is empty.
- In each transition, $A$ either pushes a symbol into the stack or pops one from the stack, but it cannot do both. More precisely, every transition can only be of the forms:

\[(p, x, \text{pop}(y)) \rightarrow (q, \text{push}(\epsilon))\]
\[(p, x, \text{push}(\epsilon)) \rightarrow (q, \text{push}(z))\]

Consider the following CFG $G = (\Sigma, V, R, S)$ where each component is as follows.

- $V = \{A_{p,q} | p, q \in Q\}$.
- $A_{q_0,q_f}$ is the start variable.
- $R$ consists of the following rules:

- For every state $p, q, r, s \in Q$ and every symbol $z \in \Gamma$ and every symbol $a, b \in \Sigma \cup \{\epsilon\}$, if the following transitions are in $\delta$:

\[(p, a, \text{pop}(\epsilon)) \rightarrow (r, \text{push}(z))\]
\[(s, b, \text{pop}(z)) \rightarrow (q, \text{push}(\epsilon))\]

then the following rule is in $R$:

\[A_{p,q} \rightarrow a A_{r,s} b \quad (R1)\]
– For every state \( p, q, r \in Q \) and every symbol \( a \in \Sigma \cup \{\epsilon\} \), if the following transition is in \( \delta \):
\[
(p, a, \text{pop}(\epsilon)) \rightarrow (r, \text{push}(\epsilon))
\]
then the following rule is in \( R \):
\[
A_{p,q} \rightarrow a A_{r,q} \quad (R2)
\]

– For every state \( p, q, r \in Q \), we have the following rule in \( R \):
\[
A_{p,q} \rightarrow A_{p,r} A_{r,q} \quad (R3)
\]

– For every \( p \in Q \), we have the following rule in \( R \):
\[
A_{p,p} \rightarrow \epsilon \quad (R4)
\]

We will show that \( L(A) = L(G) \). First, we prove the following lemma.

**Lemma 5.5** For every derivation:
\[
A_{p,q} \Rightarrow^* w, \quad \text{where } w \in \Sigma^*
\]
there is a run:
\[
(p, \epsilon) \vdash^* (q, \epsilon)
\]

**Proof.** Suppose \( A_{p,q} \Rightarrow^* w \), where \( w \in \Sigma^* \). Let the derivation of length \( m \) and denoted as follows.
\[
A_{p,q} \Rightarrow w_1 \Rightarrow \cdots \Rightarrow w_m, \quad \text{where } w_m = w. \quad (6)
\]

We will prove the lemma by induction on \( m \). The base case is \( m = 1 \). So, we have a derivation:
\[
A_{p,q} \Rightarrow w
\]
The only rule that can be used in this case is rule \((R4)\), which means \( w = \epsilon \) and \( p = q \). Thus, there is a (trivial) run:
\[
(p, \epsilon) \vdash^\epsilon (p, \epsilon)
\]

For the induction hypothesis, we assume the lemma holds for every derivation of length \( \leq m-1 \). We will prove the case of derivations of length \( m \) for the induction step.

Let the derivation be as in (6). We consider the rules applied on the step \( A_{p,q} \Rightarrow w_1 \). There are three cases:

- The rule applied is \((R1)\) type of rule, say:
\[
A_{p,q} \rightarrow a A_{r,s} b
\]
This means that \( w_1 = a A_{r,s} b \), and hence, every \( w_i \) starts with \( a \) and ends with \( b \), for every \( i \in \{1, \ldots, m\} \). We denote each \( w_i = aw'_i b \), for some \( w'_i \). Thus, there is a derivation:
\[
A_{r,s} \Rightarrow w'_2 \Rightarrow \cdots \Rightarrow w'_m
\]
This derivation has length \( m-1 \). By the induction hypothesis, there is a run:
\[
(r, \epsilon) \vdash_{w'_{m+1}}^* (s, \epsilon) \quad (7)
\]
Since \( A_{p,q} \to aA_{r,q}b \) is a rule in \( R \), there are the following two transitions in \( \delta \):
\[
(p, a, \text{pop}(\epsilon)) \to (r, \text{push}(z)) \quad \text{and} \quad (s, b, \text{pop}(z)) \to (q, \text{push}(\epsilon))
\] (8)

Now, from the run (7), since the content of the stack never goes “below empty”, we also have a run:
\[
(r, z) \vdash_{w_{m+1}'} (s, z)
\] (9)

Using the transitions (5), we have:
\[
(p, \epsilon) \vdash_a (r, z) \quad \text{and} \quad (s, z) \vdash_b (q, \epsilon)
\] (10)

Combining the runs (9) and (10), we have:
\[
(p, \epsilon) \vdash^* (r, \epsilon) \quad \text{and} \quad (r, \epsilon) \vdash^* (q, \epsilon)
\]

That is,
\[
(p, \epsilon) \vdash^*_w (r, \epsilon)
\]

Since \( w = aw_{m+1}'b \), the run is as desired.

- The rule applied is \((R.3)\) type of rule, say:
\[
A_{p,q} \to A_{p,r}A_{r,q}, \quad \text{for some } r \in Q.
\]

Thus, \( w_1 = A_{p,r}A_{r,q} \). This means \( w = xy \) where \( A_{p,r} \Rightarrow^* x \) and \( A_{r,q} \Rightarrow^* y \). Both are derivations of length \( \leq m - 1 \). By the induction hypothesis, there are the following two runs:
\[
(p, \epsilon) \vdash^*_x (r, \epsilon) \quad \text{and} \quad (r, \epsilon) \vdash^*_y (q, \epsilon)
\]

Combining these two runs, we have:
\[
(p, \epsilon) \vdash^*_w (r, \epsilon)
\]

- The rule applied is \((R.2)\) type of rule, say:
\[
A_{p,q} \to aA_{r,q}, \text{ for some } r \in Q \text{ and } a \in \Sigma
\]

Thus, \( w_1 = aA_{r,q} \). This means that every \( w_i \) starts with \( a \), for every \( i \in \{1, \ldots, m\} \). We denote each \( w_i = aw_i' \), for some \( w_i' \). Thus, there is a run:
\[
A_{r,q} \Rightarrow w_2' \Rightarrow \cdots \Rightarrow w_m'.
\]

The length of this run is \( m - 1 \). By the induction hypothesis, there is a run:
\[
(r, \epsilon) \vdash^*_w (q, \epsilon)
\] (11)

Now, since there is a rule \( A_{p,q} \to aA_{r,q} \), there is the following transition in \( \delta \):
\[
(p, a, \text{pop}(\epsilon)) \to (r, \text{push}(\epsilon))
\]
Using this transition, we have:

\[(p, \epsilon) \vdash_a (r, \epsilon)\]

Combining this with run \([11]\), we have run:

\[(p, \epsilon) \vdash_a (r, \epsilon) \vdash^{w^*}_{u_m} (q, \epsilon)\]

Thus, there is a run:

\[(p, \epsilon) \vdash^{w^*}_{u_m} (q, \epsilon)\]

This run is as required.

This completes the proof of Lemma 5.5.

The following lemma is the converse direction of the previous lemma.

**Lemma 5.6** For every run:

\[(p, \epsilon) \vdash^* w (q, \epsilon)\]

there is a derivation:

\[A_{p,q} \Rightarrow^* w\]

**Proof.** Suppose there is a run \((p, \epsilon) \vdash^* w (q, \epsilon)\). Let the run has length \(n\), denoted as follows.

\[(p_0, v_0) \vdash_{b_1} (p_1, v_1) \vdash_{b_2} \cdots \vdash_{b_n} (p_n, v_n)\] (12)

where \(p_0 = p, v_0 = \epsilon, p_n = q, v_n = \epsilon\) and \(b_1 \cdots b_n = w\). We will prove the lemma by induction on \(n\).

The base case is \(n = 0\). In this case the run is a trivial run:

\[(p_0, \epsilon) \vdash_{\epsilon} (p_0, \epsilon)\]

Applying the \((R4)\) type of rule \(A_{p_0,p_0} \rightarrow \epsilon\), we have a derivation:

\[A_{p_0,p_0} \Rightarrow \epsilon\]

For the induction hypothesis, we assume that the lemma holds for run of length \(n \leq n - 1\). We will now prove it for runs of length \(n\) for the induction step.

Consider a run of length \(n\) as in (12). We consider the transition used in the step \((p_0, v_0) \vdash_{b_1} (p_1, v_1)\). There are two cases:

- The transition used is \((p_0, b_1, \text{pop}(\epsilon)) \rightarrow (p_1, \text{push}(\epsilon)), i.e., the PDA \(A\) pops and pushes nothing to its stack. By definition of the CFG \(G\), there is a \((R2)\) type of rule:

\[A_{p_0,p_n} \rightarrow b_1 A_{p_1,p_n}\] (13)

Applying the induction hypothesis on the run:

\[(p_1, v_1) \vdash_{b_2} \cdots \vdash_{b_n} (p_n, v_n)\]

there is a derivation:

\[A_{p_1,p_n} \Rightarrow^* b_2 \cdots b_n\]

Moreover, using rule \([13]\), we have derivation:

\[A_{p_0,p_n} \Rightarrow b_1 A_{p_1,p_n} \Rightarrow^* b_1 b_2 \cdots b_n\]

as desired.
The transition used is of the form: $(p_0, b_1, \text{pop}(\epsilon)) \to (p_1, \text{push}(z))$, for some $z \in \Gamma$.

There are two more cases here.

- For some $j \in \{2, \ldots, n\}$, $v_j = \epsilon$.
  Thus, there is a run:
  $$(p_0, v_0) \vdash^* (p_j, v_j) \vdash^* (p_n, v_n)$$
  where $w = xy$.

  By induction hypothesis, there are derivations:
  $$A_{p_0, p_j} \Rightarrow^* x \quad \text{and} \quad A_{p_j, p_n} \Rightarrow^* y$$

  By construction of $G$, there is a $[R3]$ type of rule:
  $$A_{p_0, p_n} \to A_{p_0, p_j} A_{p_j, p_n}.$$

  Using this rule, we have the following derivation:
  $$A_{p_0, p_n} \Rightarrow A_{p_0, p_j} A_{p_j, p_n} \Rightarrow^* xy$$

  This is the desired derivation, since $w = xy$.

- For every $i \in \{1, \ldots, n - 1\}$, $v_i \neq \epsilon$.
  This means that $z$ is popped from $v_{n-1}$ to obtain $v_n$ and there are the following transitions:
  $$(p_0, a, \text{pop}(\epsilon)) \to (p_1, \text{push}(z)) \quad \text{and} \quad (p_{n-1}, b, \text{pop}(z)) \to (p_n, \text{push}(\epsilon))$$
  (14)

  where $b_1 = a$ and $b_n = b$. Since $v_n = \epsilon$, this means $v_{n-1} = z$. In particular, every $v_i$ starts with $z$, for every $i \in \{1, \ldots, n - 1\}$. We denote by $v_i = zv'_i$, for some $i \in \{1, \ldots, n - 1\}$. This means there is a run:
  $$(p_1, z) \vdash^* (p_{n-1}, z)$$
  where $w = aw'b$

  Since during this run, $z$ is untouched, we have a run:
  $$(p_1, \epsilon) \vdash^* (p_{n-1}, \epsilon)$$

  This run has length $\leq n - 1$. By the induction hypothesis, there is a derivation:
  $$A_{p_1, p_{n-1}} \Rightarrow^* w'$$

  Due to the transitions in (14), there is a $[R1]$ type of rule:
  $$A_{p_0, p_n} \to aA_{p_1, p_{n-1}} b$$

  Using this rule, there is a derivation:
  $$A_{p_0, p_n} \Rightarrow aA_{p_1, p_{n-1}} b \Rightarrow^* aw'b.$$

  Since $aw'b = w$, this derivation is as desired.

This completes the proof of Lemma 5.6. ■
Theorem 5.7 \( L(A) = L(G) \).

**Proof.** We first prove \( L(G) \supseteq L(A) \). Let \( w \in L(A) \). So, there is an accepting run of \( A \) on \( w \), which must start from the initial state \( q_0 \) and end with the accepting state \( q_f \):

\[
(q_0, \epsilon) \xrightarrow{\ast \ w} (q_f, \epsilon)
\]

By Lemma 5.6, we have:

\[
A_{q_0,q_f} \Rightarrow^* w.
\]

Thus, \( w \in L(G) \).

Now we prove that \( L(G) \subseteq L(A) \). Let \( w \in L(G) \). Thus,

\[
A_{q_0,q_f} \Rightarrow^* w.
\]

By Lemma 5.5, there is a run:

\[
(q_0, \epsilon) \xrightarrow{\ast \ w} (q_f, \epsilon).
\]

Since this is an accepting run, \( w \in L(A) \). This completes the proof of Theorem 5.7. \( \square \)