CSCE 355
3/18/2024

CFG $\longrightarrow$ PDA : partial proof of correctness

1

Recall: CFG $G = \langle V, \Sigma, S, P \rangle$.

Construct a PDA $P = \langle \{q\}, \Sigma, V \cup \Sigma, \delta, q, S, \emptyset \rangle$ such that $L(G) = N(P)$, where

$$\delta(q, a, a) = \{(q, \varepsilon)\} \longleftarrow \text{matching transitions } \forall a \in \Sigma$$

and for all $A \in V$, let $\{A \to \alpha_1, A \to \alpha_2, \ldots\}$ be the set of all productions with head A.

$$\delta(q, \varepsilon, A) = \{(q, \alpha_1), (q, \alpha_2), \ldots\}$$

Today: Prove that $N(P) \subseteq L(G)$.

Def: ~~Lemma~~: Let $\alpha \in V \cup \Sigma$ such that

$S \Rightarrow^* \alpha$ [α is a sentential form].

Lemma: Given an input string $w \in \Sigma^*$, if $(q, w, S) \vdash^* (q, y, \gamma)$ for some $y \in \Sigma^*$ (y is a suffix of w) and $\gamma \in (V \cup \Sigma)^*$.

Let $x \in \Sigma^*$ be such that $w = xy$
[x is uniquely determined: input symbols consumed]

Then $S \Rightarrow^* x\gamma$ ($x\gamma$ is a sentinal form).

zero or more steps in a derivation

Proof is by induction on the number n of steps in the computation path.

Base case: $n=0$. No steps, so $(q, w, S) = (q, y, \gamma)$, that is $y = w$ and $\gamma = S$.

So $x = \varepsilon$. So $x\gamma = \varepsilon S = S$, and clearly, $S \Rightarrow^* S$ (0 steps)

$\therefore$ Lemma holds for $n=0$.

Inductive case: Suppose lemma holds for some $n \geq 0$. Show the lemma holds for $n+1$. We assume a computation path

$$\underbrace{(q, w, S) \vdash \cdots \vdash (q, y_n, \gamma_n)}_{n \text{ steps}} \vdash (q, y_{n+1}, \gamma_{n+1})$$

By the inductive hypothesis, $S \Rightarrow^* x_n \gamma_n$, where x_n is unique such that $x_n y_n = w$

Consider the last step:

$\gamma_n \neq \varepsilon$. Let X be the first symbol of γ_n

<u>Case 1</u>: $X \in \Sigma$ (X is a terminal).

Then $(q, y_n, \gamma_n) \vdash (q, y_{n+1}, \gamma_{n+1})$ is a matching transition: X must be the first symbol of y_n and so

$$\underbrace{y_n = X y_{n+1}}_{X \text{ is consumed off the input}} \quad \text{and} \quad \underbrace{\gamma_n = X \gamma_{n+1}}_{X \text{ is popped off the stack}}$$

Know that ~~$x_n y_n$~~ $w = x_n y_n$ (unique $x_n \in \Sigma^*$)

and $w = x_{n+1} y_{n+1}$ (" $x_{n+1} \in \Sigma^*$)

So $w = x_n X y_{n+1} = x_{n+1} y_{n+1}$ so $x_{n+1} = x_n X$

~~Similarly~~ So $x_{n+1} \gamma_{n+1}$ (want to show this is a sentential form)

$= x_n X \gamma_{n+1} = x_n \gamma_n$ ← sentential form by the inductive hyp.

So no additional step in the derivation.

Case 2: Not Case 1. $X \in V$.

Then $(q, y_n, \gamma_n) \vdash (q, y_{n+1}, \gamma_{n+1})$ is an expansion step, in particular, an ε-move, so $y_{n+1} = y_n$ and $x_{n+1} = x_n$, and ~~γ_{n+1}~~ $\gamma_n = A\gamma$ for $A \in V$ and some $\gamma \in (V \cup \Sigma)^*$ and $\gamma_{n+1} = \alpha\gamma$, where $A \to \alpha$ is a production of G.

Then $x_{n+1}\gamma_{n+1} = x_{n+1}\alpha\gamma = x_n\alpha\gamma$.

By the inductive hyp., $S \Rightarrow^* x_n\gamma_n = x_n A\gamma$, but $x_n A\gamma \Rightarrow x_n\alpha\gamma$, so

$$S \Rightarrow \cdots \Rightarrow x_n A\gamma \Rightarrow x_n\alpha\gamma = x_{n+1}\gamma_{n+1}$$

$\therefore$ $x_{n+1}\gamma_{n+1}$ is sentl'al form.

Proves the lemma by induction on n $\blacksquare$

~~Cor~~ Theorem: $N(P) \subseteq L(G)$.

Proof: Let $w \in \Sigma^*$ be any input string to P.

$$w \in N(P) \iff (q, w, S) \vdash^* (q, \varepsilon, \varepsilon)$$

By the lemma, this says that

$$S \Rightarrow^* w\varepsilon = w$$

all of w consumed   stack contents at the end.

$\therefore$ $w \in L(G)$. Since w was arbitrary, $N(P) \subseteq L(G)$. ◪

<u>Thm</u>: $L(G) \subseteq N(P)$ [proof omitted]

<u>Corr</u>: For every CFG G there exists a 1-state PDA P such that $N(P) = L(G)$

<u>Cor</u>: - - - - - - $L(P) = L(G)$

<u>Next up</u>: PDA $\longrightarrow$ CFG

2 steps: PDA $\longrightarrow$ "Restricted PDA" $\longrightarrow$ CFG

<u>Def</u>: A <u>restricted</u> <u>PDA</u> is a PDA

$$P = \langle Q, \Sigma, \Gamma, \delta, q_0, Z_0, \emptyset \rangle$$

where $Q, \Sigma, \Gamma, q_0, Z_0$ are as with any PDA, but for any $q \in Q$, $a \in \Sigma \cup \{\varepsilon\}$, $X \in \Gamma$

the only elements of $\delta(q, a, X)$

are either (r, ε) some $r \in Q$

denoted (r, pop)

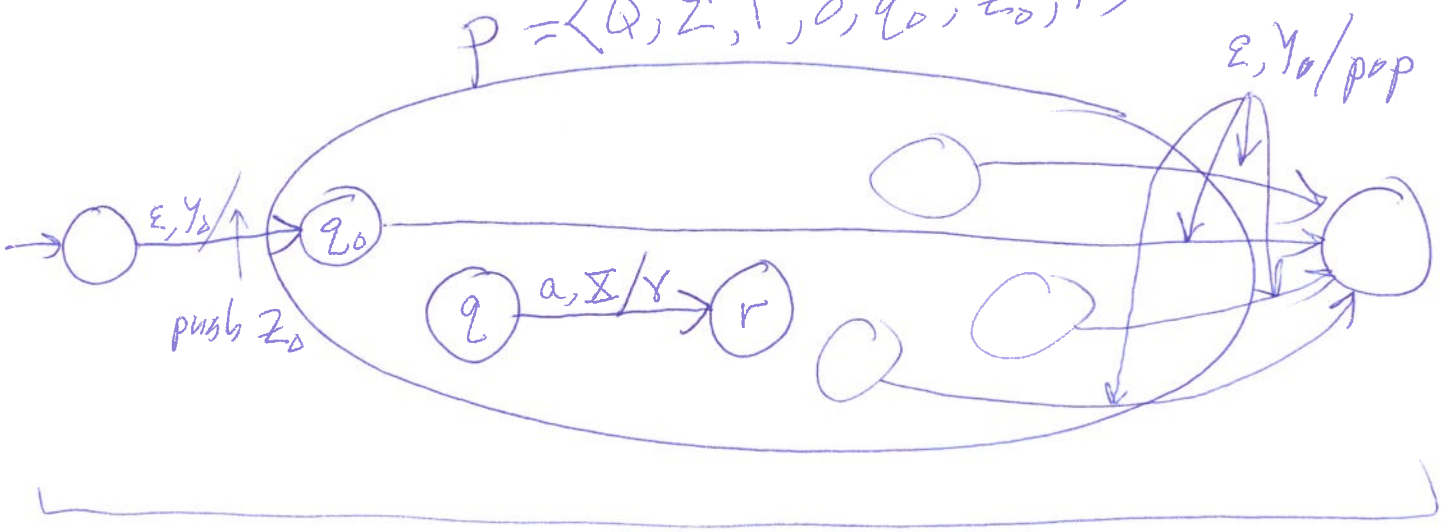
or (r, YX) for some $Y \in \Gamma$

denoted $(r, \text{push } Y)$

Prop: For any PDA P there exists an ~~equivalent~~ ~~restres~~ restricted PDA R such that $N(R) = N(P)$.

Proof: By construction:

$P = \langle Q, \Sigma, \Gamma, \delta, q_0, Z_0, F \rangle$



R

$Y_0 \notin \Gamma$ is the bottom stack marker for R

For every transition $q \xrightarrow{a, X / \gamma} r$ of P, let $\gamma := X_n X_{n-1} \cdots X_1$ where each $X_i \in \Gamma$. Replace this transition with

$$q \xrightarrow{a, X / \text{pop}} \bigcirc \xrightarrow{\varepsilon, \text{anything} / \text{push } X_1} \bigcirc \xrightarrow{\varepsilon, \text{anything} / \text{push } X_2} \bigcirc \rightarrow \cdots$$

$$\cdots \xrightarrow{\varepsilon, \text{anything} / \text{push } X_n} r$$

where the intermediate states are fresh states used for this transition only.

Proof of correctness omitted. ∎

Next time: Restricted PDA $\longrightarrow$ CFG