

Otomat đẩy xuống

Trần Vĩnh Đức



Trường Đại Học Bách Khoa Hà Nội

Ngày 27 tháng 3 năm 2019

Thuật ngữ

- ▶ Context-Free Language (CFL) : Ngôn ngữ phi ngữ cảnh
- ▶ Context-Free Grammar (CFG): Văn phạm phi ngữ cảnh
- ▶ Pushdown Automata (PDA): Otomat đẩy xuống

Nội dung

Giới thiệu

Định nghĩa hình thức

Ví dụ

Tương đương với văn phạm phi ngữ cảnh

Nội dung

Giới thiệu

Định nghĩa hình thức

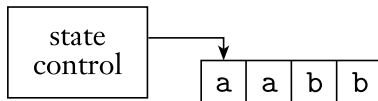
Ví dụ

Tương đương với văn phạm phi ngữ cảnh

Otomat đẩy xuống (PDA)

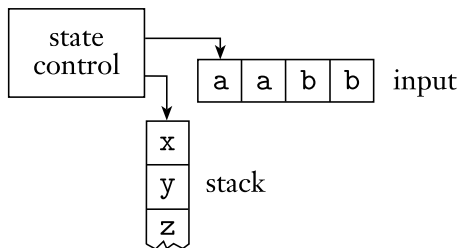
- ▶ Giống NFA nhưng có thêm một *ngăn xếp*
- ▶ Ngăn xếp cho phép PDA đoán nhận nhiều ngôn ngữ không chính quy
- ▶ Về khả năng định nghĩa ngôn ngữ, PDA tương đương với văn phạm phi ngữ cảnh

NFA ...



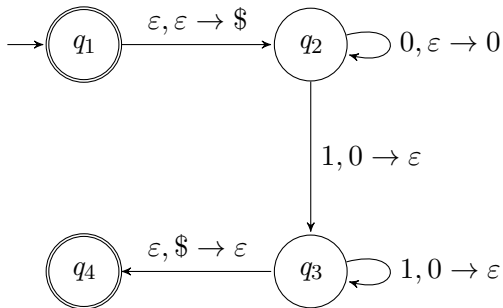
- ▶ Chỉ lưu trữ thông tin hữu hạn
- ▶ Bộ điều khiển thể hiện hàm chuyển trạng thái, băng chứa xâu vào

và PDA



- ▶ Ngăn xếp cho phép lưu trữ thông tin không giới hạn
- ▶ PDA đọc xâu vào, đọc và ghi dữ liệu của ngăn xếp, và chuyển trạng thái

Ví dụ



Hình: PDA đoán nhận ngôn ngữ $\{0^n 1^n \mid n \geq 0\}$

PDA đơn định và đa định

- ▶ Nhắc lại rằng NFA và DFA đoán nhận cùng lớp ngôn ngữ chính quy
- ▶ Khác với otomat hữu hạn, PDA đa định và đơn định không tương đương về đoán nhận ngôn ngữ
- ▶ Có nhiều ngôn ngữ đoán nhận bởi PDA đa định nhưng không đoán nhận bởi PDA đơn định nào!
- ▶ Ta sẽ chỉ xem xét các PDA đa định vì nó tương đương với CFG

Nội dung

Giới thiệu

Định nghĩa hình thức

Ví dụ

Tương đương với văn phạm phi ngữ cảnh

Bảng chữ và xâu rỗng

- ▶ Một PDA xác định hai bảng chữ *khác nhau*: Bảng chữ cho xâu vào Σ và bảng chữ cho ngăn xếp Γ
- ▶ ε là xâu rỗng trên cả hai bảng chữ
- ▶ $\varepsilon \in \Sigma^*$ cho phép chuyển trạng thái của máy mà không cần đọc xâu vào
- ▶ $\varepsilon \in \Gamma^*$ cho phép chuyển trạng thái của máy mà không cần lấy ký hiệu trên đỉnh ngăn xếp
- ▶ Đẩy ε vào ngăn xếp không làm thay đổi ngăn xếp

Hàm chuyển trạng thái

Hàm chuyển trạng thái, δ , phụ thuộc vào

1. trạng thái hiện tại
2. ký hiệu vào
3. ký hiệu trên đỉnh ngăn xếp

Dựa vào các giá trị này, máy thay đổi trạng thái và đỉnh ngăn xếp. Vì ta cho phép đa định, nên đầu ra của hàm là một tập của các cặp gồm

1. một trạng thái mới
2. ký hiệu mới trên đỉnh ngăn xếp

$$\delta : Q \times \Sigma_{\epsilon} \times \Gamma_{\epsilon} \rightarrow \mathcal{P}(Q \times \Gamma_{\epsilon})$$

Định nghĩa

Một *Otomat đẩy xuống* là một bộ sáu

$$(Q, \Sigma, \Gamma, \delta, q_0, F)$$

trong đó

1. Q là tập trạng thái,
2. Σ là bảng chữ vào,
3. Γ là bảng chữ ngăn xếp,
4. $\delta : Q \times \Sigma_\epsilon \times \Gamma_\epsilon \rightarrow \mathcal{P}(Q \times \Gamma_\epsilon)$ là hàm chuyển trạng thái,
5. q_0 là trạng thái bắt đầu, và
6. $F \subseteq Q$ là tập trạng thái kết thúc.

Tính toán của PDA

Một PDA $M = (Q, \Sigma, \Gamma, \delta, q_0, F)$ chấp nhận chuỗi w nếu

- ▶ w có thể viết dưới dạng $w = w_1 w_2 \dots w_m$, với $w_i \in \Sigma_\epsilon$,
- ▶ và tồn tại một dãy trạng thái $r_0, r_1, \dots, r_m \in Q$
- ▶ và tồn tại các chuỗi $s_0, s_1, \dots, s_m \in \Gamma^*$

thỏa mãn ba điều kiện sau:

1. $r_0 = q_0$ và $s_0 = \epsilon$.
2. Với $i = 0, \dots, m$, ta có

$$(r_{i+1}, b) \in \delta(r_i, w_{i+1}, a)$$

với $s_i = at$ và $s_{i+1} = bt$ với $a, b \in \Gamma_\epsilon$ và $t \in \Gamma^*$.

3. $r_m \in F$.

Nội dung

Giới thiệu

Định nghĩa hình thức

Ví dụ

Tương đương với văn phạm phi ngữ cảnh

Ví dụ 1: Đoán nhận ngôn ngữ $\{0^n 1^n \mid n \geq 0\}$

Xét PDA $M_1 = (Q, \Sigma, \Gamma, \delta, q_1, F)$ với

$$Q = \{q_1, q_2, q_3, q_4\},$$

$$\Sigma = \{0, 1\},$$

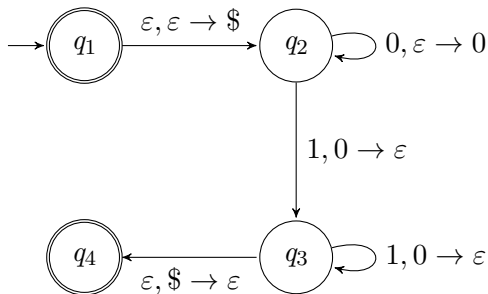
$$\Gamma = \{0, \$\},$$

$$F = \{q_1, q_4\}, \text{ và}$$

δ được cho bởi bảng sau, trong đó các ô trống ký hiệu \emptyset

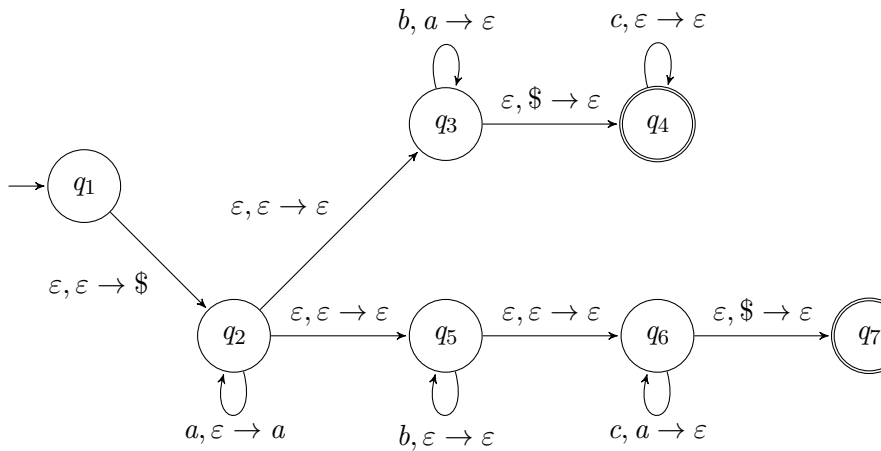
Input:	0			1			ϵ		
Stack:	0	\$	ϵ	0	\$	ϵ	0	\$	ϵ
q_1									$\{(q_2, \$)\}$
q_2			$\{(q_2, 0)\}$		$\{(q_3, \epsilon)\}$				
q_3					$\{(q_3, \epsilon)\}$			$\{(q_4, \epsilon)\}$	
q_4									

Ví dụ 1: Đoán nhận ngôn ngữ $\{0^n 1^n \mid n \geq 0\}$

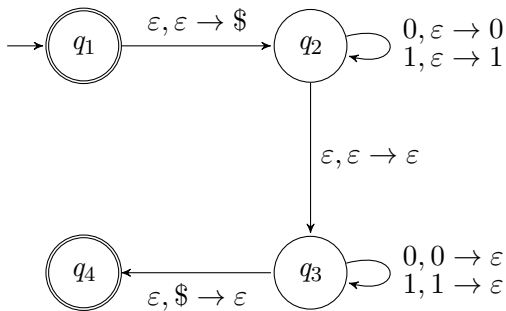


Nhãn " $a, b \rightarrow c$ ": máy đọc ký hiệu vào a và thay thế ký hiệu b trên đỉnh ngăn xếp bằng ký hiệu c .

Ví dụ 2: Đoán nhận ngôn ngữ $\{a^i b^j c^k \mid i = j \text{ hoặc } i = k\}$



Ví dụ 3: Đoán nhận ngôn ngữ $\{ww^R \mid w \in \{0, 1\}^*\}$



Nội dung

Giới thiệu

Định nghĩa hình thức

Ví dụ

Tương đương với văn phạm phi ngữ cảnh

Bổ đề

Mọi ngôn ngữ phi ngữ cảnh đều đoán nhận bởi otomat đẩy xuống.

Xây dựng PDF từ CFG

- Để thuận tiện, ta sử dụng thao tác sau trên PDA

$$(r, u_1 u_2 \dots u_\ell) \in \delta(q, a, s)$$

có nghĩa rằng máy đang ở trạng thái q , đọc ký hiệu vào a sẽ xóa ký hiệu $s \in \Gamma_\epsilon$ trên đỉnh ngăn xếp, đẩy toàn bộ xâu $u_1 u_2 \dots u_\ell$ vào ngăn xếp, và chuyển sang trạng thái r .

- Thao tác trên được cài đặt như sau: Đưa vào các trạng thái mới $q_1, q_2, \dots, q_{\ell-1}$ và các chuyển sau

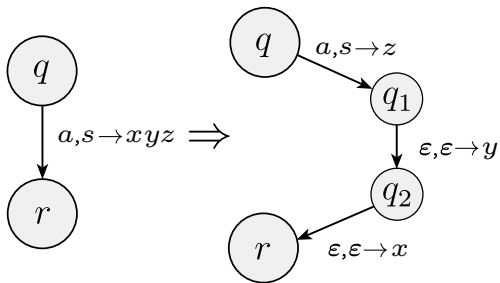
$$\delta(q, a, s) \text{ chứa } (q_1, u_\ell)$$

$$\delta(q_1, \epsilon, \epsilon) = \{(q_2, u_{\ell-1})\},$$

$$\delta(q_2, \epsilon, \epsilon) = \{(q_3, u_{\ell-2})\},$$

$$\vdots$$

$$\delta(q_{\ell-1}, \epsilon, \epsilon) = \{(r, u_1)\}.$$



Hình: Cài đặt thao tác $(r, xyz) \in \delta(q, a, s)$

Xây dựng PDA từ CFG

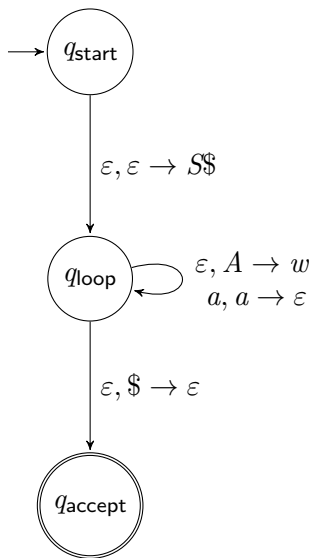
- ▶ Ta xây dựng PDA $P = (Q, \Sigma, \Gamma, \delta, q_{\text{start}}, \{q_{\text{accept}}\})$ tương đương với CFG G như sau.
- ▶ Tập trạng thái $Q = \{q_{\text{start}}, q_{\text{loop}}, q_{\text{accept}}\} \cup E$ với E là các trạng thái phát sinh trong quá trình cài đặt thao tác vừa mô tả ở trên.
- ▶ Hàm δ được định nghĩa bởi:

$$\delta(q_{\text{start}}, \varepsilon, \varepsilon) = \{(q_{\text{loop}}, S\$)\},$$

$$\delta(q_{\text{loop}}, \varepsilon, A) = \{(q_{\text{loop}}, w) \mid A \rightarrow w \text{ là một quy tắc của } G\},$$

$$\delta(q_{\text{loop}}, a, a) = \{(q_{\text{loop}}, \varepsilon)\},$$

$$\delta(q_{\text{loop}}, \varepsilon, \$) = \{(q_{\text{accept}}, \varepsilon)\}.$$



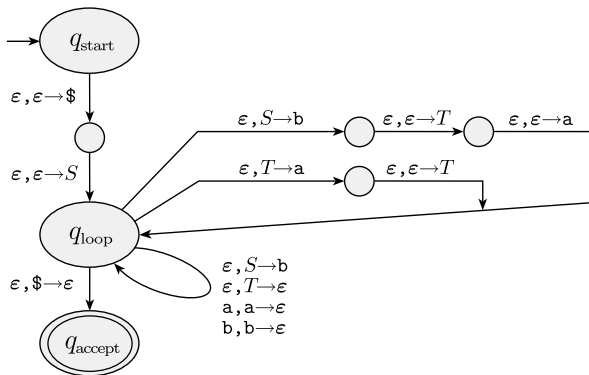
với quy tắc $A \rightarrow w$
ký hiệu kết thúc a

Ví dụ

Xét văn
phạm

$$S \rightarrow aTb \mid b$$

$$T \rightarrow Ta \mid \varepsilon$$



Bổ đề

Mọi ngôn ngữ đoán nhận bởi otomat đẩy xuống đều là phi ngữ cảnh.

CFG \leftrightarrow PDA

Định lý

Một ngôn ngữ là phi ngữ cảnh nếu và chỉ nếu có một otomat đẩy xuống đoán nhận nó.

Vì mọi NFA đều là PDA nên ta có:

Hệ quả

Mọi ngôn ngữ chính quy đều là phi ngữ cảnh.