Tạp chí Tin học và Điều khiển học, T. 15, S. 2 (1999), 20–26

CẬN DƯỚI ĐỘ PHỨC TẠP OTOMAT HỮU HẠN ĐOÁN NHẬN SIÊU NGÔN NGỮ CHÍNH QUY

ĐẶNG HUY RUẬN, PHÙNG VĂN ỔN

Abstract. Some results of regular hyper-language [1] and the complexity of finite automaton recognizing regular hyper-language [2] are considered.

In this paper we consider the lower limits of the complexity of finite automaton recognizing regular hyper-language.

1. ĐẶT VẤN ĐỀ

Nghiên cứu về lớp ngôn ngữ chính quy với từ vô hạn, chúng tôi đã thu được một số kết quả về tính chất của chúng [1]. Đặc biệt khi xem xét về độ phức tạp otomat hữu hạn đoán nhận siêu ngôn ngữ chính quy, chúng tôi nhận được đánh giá sau [2]

$$P(L) \leq 2^{2^{\cdot^{\cdot^{2^{|L|}}}}}\Big\}_{d(L)+1 \text{ lần}} + 1.$$

Vấn đề đặt ra là liệu có thể xây dựng được siêu ngôn ngữ chính quy mà độ phức tạp otomat hữu hạn đoán nhận nó đạt tới giá trị trên hay không.

Để giải quyết vấn đề trên chúng tôi dựa trên biểu thức siêu chính quy.

2. ĐÁNH GIÁ CẬN DƯỚI ĐỘ PHỨC TẠP OTOMAT HỮU HẠN ĐOÁN NHẬN SIÊU NGÔN NGỮ CHÍNH QUY

2.1. Biểu thức siêu chính quy

Giả sử A, B là biểu thức chính quy và tương ứng với chúng là ngôn ngữ chính quy. Ta định nghĩa biểu thức siêu chính quy và tương ứng với chúng là siêu ngôn ngữ chính quy bằng quy nạp như sau:

a) Siêu lặp của A - kí hiệu là $A^\infty = \{x_1x_2\ldots \in \Sigma^\alpha \mid x_i \in A\}$ - là biểu thức siêu chính quy.

b) Các biểu thức $A.B^\infty$, $A^\infty \vee B^\infty$, $A^\infty \wedge B^\infty$, $C(A^\infty)$, A^∞/B (phép cắt đoạn đầu) là những biểu thức siêu chính quy.

c) Chỉ những biểu thức thỏa mãn a), b) mới là biểu thức siêu chính quy.

Ngôn ngữ được sinh bởi biểu thức siêu chính quy D cũng là siêu ngôn ngữ chính quy và được kí hiệu bởi $L(D)$.

2.2. Độ dài của biểu thức siêu chính quy [3]

Giả sử A, B là các biểu thức chính quy và $\beta(A), \beta(B)$ là độ dài của chúng, khi đó độ dài của các biểu thức siêu chính quy tương ứng được xác định như sau: $\beta(A^\infty) = \beta(A)$, $\beta(A.B^\infty) = \beta(A) + \beta(B)$, $\beta(A^\infty \vee B^\infty) = \beta(A) + \beta(B)$, $\beta(A^\infty \wedge B^\infty) = \beta(A) + \beta(B)$, $\beta(C(A^\infty)) = \beta(A)$.

2.3. Định lý 1. *Đối với mỗi số tự nhiên t có thể xây dựng được biểu thức siêu chính quy $A_{t,n}$ trên bảng gồm 3 chữ cái như sau:*

a) Với mỗi số tự nhiên n, biểu thức siêu chính quy $A_{t,n}$ chứa t ký hiệu phép lấy phần bù và có độ dài không lớn hơn n.

b) Với mỗi số thực $c > 2$, với số tự nhiên n đủ lớn thì $\underbrace{\log_2\log_2\ldots\log_2}_{t+1 \text{ lần}} P(A_{t,n}) \geq \dfrac{n}{c\log_2 n}$.

Chứng minh.

1. Đặt $k = t + 1$. Lập bảng chữ cái $\Sigma = \{0, 1, x, \beta_0, \beta_1, ..., \beta_k\}$. Chọn số tự nhiên tùy ý m_0 và xây dựng tập D_0 bao gồm $2m_0$ từ trên bảng chữ cái $\{0, 1\}$. Tiếp tục xây dựng các tập $D_1, D_2, ..., D_k$ bằng quy nạp như sau: giả sử D_i đã được xây dựng với $2m_i$ phần tử, khi đó các phần tử của tập D_{i+1} sẽ bao gồm tất cả các tập con chứa m_i phần tử của tập D_i. Tập D_{i+1} sẽ có $C_{2m_i}^{m_i}$ phần tử và đây là số chẵn (vì $C_{2m_i}^{m_i} = 2C_{2m_i-1}^{m_i}$) nên ta đặt $m_{i+1} = \frac{1}{2} C_{2m_i}^{m_i}$. Ta sẽ dùng các chữ cái l^i, h^i (có hoặc không có chỉ số) để kí hiệu các phần tử của tập D_i.

2. Ta xây dựng các biểu thức sau:

$$\mathcal{R}_i = (0 \vee 1 \vee x \vee \beta_0 \vee \beta_1 \vee \cdots \vee \beta_i)^*$$
$$\mathcal{E}_i = \beta_i \vee \beta_{i+1} \vee \cdots \vee \beta_k \qquad 0 \le i \le k$$
$$\Sigma^* = (0 \vee 1 \vee x \vee \mathcal{E}_0)^*$$
$$\mathcal{B}_0 = \bigcup_{l^0 \in D_0} l^0\, \mathcal{E}_0\, \Sigma^*\, \mathcal{E}_0\, l^0$$
$$\mathcal{B}_{i+1} = C(\mathcal{R}_i\, \beta_i\, \mathcal{B}_i\, \beta_i\, \mathcal{R}_i) \qquad 0 \le i < k$$

Dễ nhận thấy $\mathcal{R}_i$, $\mathcal{E}_i$, Σ^*, $\mathcal{B}_0$, $\mathcal{B}_i$ là các biểu thức chính quy. Ta có thể chứng minh được độ phức tạp otomat hữu hạn đoán nhận ngôn ngữ chính quy sinh bởi $\mathcal{B}_k$ không nhỏ hơn $2m_k$ (tức $\mathcal{P}(\mathcal{B}_k) \ge 2m_k$) như sau:

• Đối với mỗi phần tử $l^i \in D_i$ ta định nghĩa các từ $\lambda_i(l^i)$ và $\overline{\lambda}_i(l^i)$ bằng quy nạp như sau:

a) $\lambda_0(l^0)$ và $\overline{\lambda}_0(l^0) = l^0$.

b) Giả sử l^{i+1} đã xác định. Với mọi $l^i \in l^{i+1}$ và từ các từ $\beta_i\lambda_i(l^i)$ ta xây dựng từ $\lambda_{i+1}(l^{i+1})$ bằng cách ghép các từ $\beta_i\lambda_i(l^i)$ theo một trình tự nào đó (trình tự từ điển chẳng hạn). Tương tự, với mọi $l^i \notin l^{i+1}$ và từ các từ $\overline{\lambda}_i(l^i)\beta_i$ ta xây dựng từ $\overline{\lambda}_{i+1}(l^{i+1})$ bằng cách ghép các từ $\overline{\lambda}_i(l^i)\beta_i$ theo một trình tự nào đó. Dễ nhận thấy $\lambda_{i+1}(l^{i+1}) \in \mathcal{R}_i$, $\overline{\lambda}_{i+1}(l^{i+1}) \in \mathcal{R}_i$.

Bổ đề 1. *Nếu $\lambda_{i+1}(l^{i+1})\beta = X\beta_i Y$ thì có thể tìm được l^i và β' để $\lambda_i(l^i)\beta'$ là đoạn đầu của Y trong đó $\beta \in \mathcal{E}_{i+1}$, $\beta' \in \mathcal{E}_i$, $l^i \in l^{i+1}$.*

Thật vậy, giả sử $\lambda_{i+1}(l^{i+1})\beta = X\beta_i Y$, theo định nghĩa $\lambda_{i+1}(l^{i+1})$ được tạo thành bằng cách ghép các từ dạng $\beta_i\lambda_i(l^i)$ lại với nhau theo một trình tự nào đó (ở đây $l^i \in l^{i+1}$); như vậy trong từ $\lambda_i(l^i)$ không chứa ký hiệu β_i. (Điều này giải thích lý do đưa kí hiệu β_i vào vế phải đẳng thức trên để tách từ đó ra thành 2 nửa). Do vậy nhất định sẽ xảy ra hoặc $\beta_i Y = \beta_i\lambda_i(l^i)\beta$ hoặc $\beta_i Y$ phải được bắt đầu bằng từ $\beta_i\lambda_i(l^i)\beta_i$.

Ta nhận được từ khả năng thứ nhất là $\beta' = \beta$ còn từ khả năng thứ hai là $\beta' = \beta_i$. Vậy $\lambda_i(l^i)\beta'$ là đoạn đầu của Y.

Một cách tương tự, nếu $\beta\overline{\lambda}_{i+1}(l^{i+1}) = X\beta_i Y$ thì có thể tìm được l^i và β' để $\lambda_i(l^i)\beta'$ là đoạn cuối của X trong đó $\beta \in \mathcal{E}_{i+1}$, $\beta' \in \mathcal{E}_i$, $l^i \notin l^{i+1}$.

Bổ đề 2. *Nếu $l^i \in l^{i+1}$ và $\beta \in \mathcal{E}_{i+1}$ thì với β' nào đó ($\beta' \in \mathcal{E}_i$) từ $\beta_i\lambda_i(l^i)\beta'$ là một từ con của $\lambda_{i+1}(l^{i+1})\beta$.*

Thật vậy, do $\lambda_{i+1}(l^{i+1})$ được tạo thành bằng cách ghép nối tất cả các từ $\beta_i\lambda_i(l^i)$ mà $l^i \in l^{i+1}$, vì vậy nếu $l^i \in l^{i+1}$ thì từ $\lambda_{i+1}(l^{i+1})$ phải chứa từ $\beta_i\lambda_i(l^i)$. Nếu $\beta_i\lambda_i(l^i)$ là đoạn cuối của $\lambda_{i+1}(l^{i+1})$ thì buộc $\beta = \beta'$, dẫn tới $\lambda_{i+1}(l^{i+1})\beta$ kết thúc bằng từ $\beta_i\lambda_i(l^i)$. Trường hợp ngược lại, trong từ $\lambda_{i+1}(l^{i+1})$, sau từ $\beta_i\lambda_i(l^i)$ phải là β_i ($\beta_i \in \mathcal{E}_i$), như vậy từ $\beta_i\lambda_i(l^i)\beta_i$ nằm trong $\lambda_{i+1}(l^{i+1})\beta$.

Tương tự ta có $\beta'\overline{\lambda}_i(l^i)\beta_i$ là một từ con của $\beta\overline{\lambda}_{i+1}(l^{i+1})$.

• Đối với mỗi phần tử $l^i \in D_i$ ta xây dựng các tập $K_i(l^i)$ bằng quy nạp như sau:

$$K_0(l^0) = \Sigma^*\mathcal{E}_0 l^0,$$
$$K_{i+1}(l^{i+1}) = C\Big(\big(\bigcup_{l^i \in l^{i+1}} K_i(l^i)\big)\beta_i\mathcal{R}_i\Big), \quad 0 \le i < k. \qquad (1)$$

Bổ đề 3. *Với bất kỳ từ X; l^i, h^i tùy ý $(0 \le i \le k)$ và s, $(s \ge i)$:*

$$X\beta_s\overline{\lambda}_i(h^i) \in K_i(l^i) \equiv l^i = h^i. \quad (2)$$

Chứng minh quy nạp theo i:

Giả sử $i = 0$, khi đó $K_0(l^0) = \Sigma^*\mathcal{E}_0 l^0$ và $X\beta_s\overline{\lambda}_0(h^0) = X\beta_s h^0$. Ta có $X \in \Sigma^*$, $\beta_s \in \mathcal{E}_0$ vì vậy nếu $l^0 = h^0$ thì $X\beta_s\overline{\lambda}_0(h^0) \in K_0(l^0) = \Sigma^*\mathcal{E}_0 l^0$.

Ngược lại, từ tùy ý thuộc $\Sigma^*\mathcal{E}_0 l^0$ kết thúc bằng từ $\beta_r l^0$ $(0 \le r \le k)$ thì từ $X\beta_s h^0$ thuộc $\Sigma^*\mathcal{E}_0 l^0$ chỉ trong trường hợp $l^0 = h^0$.

Giả sử với giá trị i nào đó bổ đề đã được chứng minh với mọi X, l^i, h^i và β_q $(k \ge q \ge i)$. Chọn một cách tùy ý X và các l^{i+1}, h^{i+1} và β_s $(s \ge i+1)$. Giả sử rằng

$$X\beta_s\overline{\lambda}_{i+1}(h^{i+1}) \notin K_{i+1}(l^{i+1}). \quad (3)$$

Khi đó từ (1) ta có:

$$X\beta_s\overline{\lambda}_{i+1}(h^{i+1}) \in \Big(\bigcup_{l^i \in l^{i+1}} K_i(l^i)\Big)\beta_i\mathcal{R}_i. \quad (4)$$

Dẫn tới tồn tại Y, Z, và l^i để $X\beta_s\overline{\lambda}_{i+1}(h^{i+1}) = Y\beta_i Z$, $l^i \in l^{i+1}$, $Y \in K_i(l^i)$ và $Z \in \mathcal{R}_i$.

Từ Z không thể kết thúc bằng $\beta_s\overline{\lambda}_{i+1}(h^{i+1})$ bởi vì $\mathcal{R}_i$ không chứa từ có chứa β_s, nghĩa là $\beta_s\overline{\lambda}_{i+1}(h^{i+1})$ kết thúc bằng từ $\beta_i Z$, tức là $\beta_s\overline{\lambda}_{i+1}(h^{i+1}) = U\beta_i Z$.

Theo Bổ đề 1, có thể tìm được h^i và β' $(\beta' \in \mathcal{E}_i, h^i \notin h^{i+1})$ sao cho $U = V\beta'\overline{\lambda}_i(h^i)$ với V nào đó. Khi đó $Y\beta_i Z = Z\beta_s\overline{\lambda}_{i+1}(h^{i+1}) = XU\beta_i Z$, dẫn tới $Y = YU = XV\beta'\overline{\lambda}_i(h^i)$.

Theo giả thiết quy nạp, từ $XV\beta'\overline{\lambda}_i(h^i) \in K_i(l^i)$ dẫn tới $h^i = l^i$ nghĩa là $h^i \in l^{i+1}$.

Mặt khác do $h^i \notin h^{i+1}$ nên: $h^{i+1} \ne l^{i+1}$.

Giả sử ngược lại: $\qquad h^{i+1} \ne h^{i+1}$. (5)

Tập hợp h^{i+1}, l^{i+1} thuộc D_{i+1} và theo định nghĩa, chúng có m_i phần tử.

Theo (5) ta có thể tìm được h^i mà $h^i \notin h^{i+1}$ và $h^i \in l^{i+1}$. (6)

Khi đó theo Bổ đề 2, với β' nào đó $(\beta' \in \mathcal{E}_i)$, từ $\beta'\overline{\lambda}_i(h^i)\beta_i$ là từ con của $\beta_s\overline{\lambda}_{i+1}(h^{i+1})$, nghĩa là

$$\beta_s\overline{\lambda}_{i+1}(h^{i+1}) = V\beta'\overline{\lambda}_i(h^i)\beta_i Z \quad (7)$$

Như vậy Z sẽ là đoạn cuối của từ $\overline{\lambda}_{i+1}(h^{i+1})$, và như vậy $Z \in \mathcal{R}_i$. (8)

Theo giả thiết quy nạp ta có $XV\beta'\overline{\lambda}_i(h^i) \in K_i(l^i) \equiv h^i = l^i$, vì vậy từ (6) ta có:

$$XV\beta'\overline{\lambda}_i(h^i) \in \bigcup_{l^i \in l^{i+1}} K_i(l^i). \quad (9)$$

Từ (8), (9) dẫn đến $XV\beta'\overline{\lambda}_i(h^i) \in \big(\bigcup_{l^i \in l^{i+1}} K_i(l^i)\big)\beta_i\mathcal{R}_i$ có nghĩa là từ (7) dẫn tới (4) được thỏa mãn, và như vậy (3) cũng được thỏa mãn. Như vậy, ta đã chứng minh đối với từ X tùy ý, l^{i+1}, h^{i+1} tùy ý và s $(s \ge i+1)$:

$$X\beta_s\overline{\lambda}_{i+1}(h^{i+1}) \notin K_{i+1}(l^{i+1}) \equiv l^{i+1} \ne h^{i+1}.$$

Dẫn tới

$$X\beta_s\overline{\lambda}_{i+1}(h^{i+1}) \in K_{i+1}(l^{i+1}) \equiv l^{i+1} = h^{i+1}.$$

Bổ đề được chứng minh.

Hệ quả. *Tồn tại $2m_k$ tập phân biệt dạng $K_k(l^k)$.*

Nói cách khác tồn tại $2m_k$ tập phân biệt l^k và tương ứng với nó là $2m_k$ tập $K_k(l^k)$ không giao nhau từng đôi một.

Bổ đề 4. *Với bất kỳ Y và l^i tùy ý $(0 \le i \le k)$, nếu $Y \in \lambda_i(l^i)\mathcal{E}_i\Sigma^*$ thì $B_i/Y\mathcal{X} = K_i(l^i)$.*

Chứng minh quy nạp theo i:

+ Với $i = 0$, khi đó $\mathcal{B}_0 = \bigcup_{h^0 \in D_0} h^0 \mathcal{E}_0 \Sigma^* \mathcal{E}_0 h^0$.

Giả sử $Y \in \lambda_0(l^0)\mathcal{E}_0\Sigma^* = l^0\mathcal{E}_0\Sigma^*$, khi đó $Y = l^0\beta_s X$ với $X \in \Sigma^*$ và s nào đó.

Xét tập hợp $T(h^0) = h^0\mathcal{E}_0\Sigma^*\mathcal{E}_0 h^0/Y\mathcal{X}$, ta có

$$\begin{aligned} Z \in T(h^0) &\equiv Y\mathcal{X}Z \in h^0\mathcal{E}_0\Sigma^*\mathcal{E}_0h^0 \\ &\equiv l^0\beta_s X\mathcal{X}Z \in h^0\mathcal{E}_0\Sigma^*\mathcal{E}_0h^0 \\ &\equiv (l^0 = h^0 \ \&\ X\mathcal{X}Z \in \Sigma^*\mathcal{E}_0h^0). \end{aligned}$$

Như vậy, với $l^0 \neq h^0$ thì $T(h^0) \neq \emptyset$.

Trong trường hợp $l^0 = h^0$ thì tập $\mathcal{E}_0h^0$ không chứa từ mà trong nó có chữ cái $\mathcal{X}$, dẫn đến:

$$\begin{aligned} Z \in T(h^0) &\equiv Y\mathcal{X}Z \in \Sigma^*\mathcal{E}_0h^0 \\ &\equiv \exists\beta\ (\beta \in \mathcal{E}_0\ \&\ Z \text{ là phần đầu của } \beta h^0) \\ &\equiv Z \in \Sigma^*\mathcal{E}_0h^0 \equiv Z \in \Sigma^*\mathcal{E}_0l^0. \end{aligned}$$

Như vậy, $T(h^0) = \Sigma^*\mathcal{E}_0l^0$ và ta có

$$\mathcal{B}_0/Y\mathcal{X} = \bigcup_{h^0 \in D_0} (h^0\mathcal{E}_0\Sigma^*\mathcal{E}_0h^0/Y\mathcal{X}) = \bigcup_{h^0 \in D_0} T(h^0) = \Sigma^*\mathcal{E}_0l^0 = K_0(l^0).$$

+ Giả sử bổ đề đúng với từ Z bất kỳ và l^i tùy ý $(0 \leq i \leq k)$. Ta phải chứng minh với từ Y bất kỳ và l^{i+1} tùy ý, nếu $Y \in \lambda_{i+1}(l^{i+1})\mathcal{E}_{i+1}\Sigma^*$ thì $\mathcal{B}_{i+1}/Y\mathcal{X} = K_{i+1}(l^{i+1})$.

Giả sử $Y \in \lambda_{i+1}(l^{i+1})\mathcal{E}_{i+1}\Sigma^*$, khi đó $Y = \lambda_{i+1}(l^{i+1})\beta X$ với $\beta \in \mathcal{E}_{i+1}$. Xét tập $\mathcal{Y} = Y/\mathcal{R}_i\beta_i$. Nếu $Z \in \mathcal{Y}$, khi đó với U bất kỳ thuộc $\mathcal{R}_i$ ta có $U\beta_iZ = Y = \lambda_{i+1}(l^{i+1})\beta X$. Nhận thấy β không thể nằm trong U, dẫn tới $U\beta_i$ là đoạn đầu của $\lambda_{i+1}(l^{i+1})$, $\lambda_{i+1}(l^{i+1})\beta = U\beta_iV$ với V nào đó. Theo Bổ đề 1, có thể tìm được l^i và β' $(\beta' \in \mathcal{E}_{i+1})$ sao cho $l^i \in l^{i+1}$ và $\lambda_i(l^i)\beta'$ là đoạn đầu của V. Ta có $U\beta_iZ = \lambda_{i+1}(l^{i+1})\beta X = U\beta_iVX$; $Z = VX$; $\lambda_i(l^i)\beta'$ là đoạn đầu của Z và $Z \in \lambda_i(l^i)\mathcal{E}_i\Sigma^*$ và theo giả thiết quy nạp thì $\mathcal{B}_i/Z\mathcal{X} = K_i(l^i)$.

Như vậy

$$\forall Z\ (Z \in \mathcal{Y} \supset \exists l^i\ (l^i \in l^{i+1}\ \&\ \mathcal{B}_i/Z\mathcal{X} = K_i(l^i))). \tag{10}$$

Lấy l^i tùy ý $(l^i \in l^{i+1})$, khi đó theo Bổ đề 2, tồn tại β' $(\beta' \in \mathcal{E}_{i+1})$ thỏa mãn $\beta_i\lambda_i(l^i)\beta'$ nằm trong $\lambda_{i+1}(l^{i+1})\beta$. Vì vậy, tồn tại P và Q thỏa mãn $\lambda_{i+1}(l^{i+1})\beta = P\beta_i\lambda_i(l^i)\beta'Q$.

Khi đó $Y = \lambda_{i+1}(l^{i+1})\beta X = P\beta_i\lambda_i(l^i)\beta'QX$. Ta nhận được $Z = \lambda_i(l^i)\beta'QX$ và $P \in \mathcal{R}_i$.

Theo giả thiết quy nạp $\mathcal{B}_i/Z\mathcal{X} = K_i(l^i)$, dẫn đến:

$$\forall l^i\ (l^i \in l^{i+1} \supset \exists Z\ (Z \in \mathcal{Y}\ \&\ \mathcal{B}_i/Z\mathcal{X} = K_i(l^i))). \tag{11}$$

Từ (10), (11) dẫn đến

$$\bigcup_{Z \in \mathcal{Y}} \mathcal{B}_i/Z\mathcal{X} = \bigcup_{l^i \in l^{i+1}} K_i(l^i). \tag{12}$$

Ta có $\mathcal{B}_{i+1}/Y\mathcal{X} = C(\mathcal{R}_i\beta_i\mathcal{B}_i\beta_i\mathcal{R}_i)/Y\mathcal{X} = C((\mathcal{R}_i\beta_i\mathcal{B}_i\beta_i\mathcal{R}_i)/Y\mathcal{X}) =$
$C((\mathcal{R}_i\beta_i/Y\mathcal{X})\mathcal{B}_i\beta_i\mathcal{R}_i \cup \bigcup_{z \in Y/\mathcal{R}_i\beta_i} ((\mathcal{B}_i/Z\mathcal{X})\beta_i\mathcal{R}_i \cup \bigcup_{W \in Z/\mathcal{B}_i} \beta_i\mathcal{R}_i/W\mathcal{X}))$.

Do $\mathcal{R}_i$ không chứa từ mà chứa trong nó chữ cái $\mathcal{X}$, vì vậy $\mathcal{R}_i\beta_i/Y\mathcal{X} = \beta_i\mathcal{R}/W\mathcal{X} = \emptyset$, và

$$\mathcal{B}_{i+1}/Y\mathcal{X} = C\Big(\big(\bigcup_{Z \in Y/\mathcal{R}_i\beta_i} \mathcal{B}_i/Z\mathcal{X}\big)\beta_i\mathcal{R}_i\Big).$$

Từ (12)

$$\mathcal{B}_{i+1}/Y\mathcal{X} = C\Big(\big(\bigcup_{l^i \in l^{i+1}} K_i(l^i)\big)\beta_i\mathcal{R}_i\Big) = K_{i+1}(l^{i+1}).$$

Bổ đề được chứng minh.

Hệ quả. *Từ Bổ đề 3 và Bổ đề 4 ta có* $\mathcal{P}(\mathcal{B}_k) \geq 2m_k$.

Kết quả này nhận được nhờ áp dụng Định lý Myhill-Nerode trên $\mathcal{B}_k$ với phép cắt đoạn đầu của các từ và các tập $K_k(l^k)$ không giao nhau.

3. Ta thiết lập biểu thức chính quy $\mathcal{A}_t = \mathcal{R}_t\beta_t\mathcal{B}_t\beta_t\mathcal{R}_t$. Biểu thức này được xây dựng trên bảng gồm $t+5$ chữ cái.

Với $k = t+1$, ta có $\mathcal{B}_k = C(\mathcal{A}_t)$ và $\mathcal{P}(\mathcal{B}_k) = \mathcal{P}(\mathcal{A}_t) \geq 2m_k$.

4. Ta xây dựng bảng gồm 3 chữ cái $\Sigma_1 = \{0, 1, \alpha\}$. Lập ánh xạ từ Σ sang Σ_1 như sau: $\varphi(0) = 0$; $\varphi_1 = 1$; $\varphi(x) = \alpha 0\alpha$; $\varphi(\beta_i) = \alpha 0^{i+2}\alpha$ $(0 \leq i \leq k)$ và với từ tùy ý $x_1x_2\ldots x_n$ trên bảng chữ cái Σ thì $\varphi(x_1x_2\ldots x_n) = \varphi(x_1)\varphi(x_2)\ldots\varphi(x_n)$. Ta xây dựng các tập tương ứng với $\mathcal{R}_i$, $\mathcal{E}_0$, $\mathcal{B}_0$, $\mathcal{B}_{i+1}$, $\mathcal{A}_t$ trên Σ_1 như sau:

$$\mathcal{E}_i' = \bigcup_{i=0}^{k} \alpha 0^{i+2}\alpha \qquad 0 \leq i \leq k$$

$$\Sigma_1^* = (0 \vee 1 \vee \mathcal{E}_0')^*$$

$$\mathcal{R}_i' = (0 \vee 1 \vee \bigcup_{r=0}^{i} \alpha 0^{r+2}\alpha)^* \quad 0 \leq i \leq k$$

$$\mathcal{B}_0' = \bigcup_{l^0 \in D_0} l^0\mathcal{E}_0'\Sigma_1^*\mathcal{E}_0' l^0;$$

$$\mathcal{B}_{i+1}' = C(\mathcal{R}_i'\alpha 0^{i+2}\alpha\mathcal{B}_i'\alpha 0^{i+2}\alpha\mathcal{R}_i') \quad 0 \leq i < k.$$

$$\mathcal{A}_t' = \mathcal{R}_t'\alpha 0^{t+2}\alpha\mathcal{B}_t'\alpha 0^{t+2}\alpha\mathcal{R}_t'.$$

Trong [3] đã chứng minh được có thể lấy $c > 2$, độ dài của $\mathcal{A}_t$ nhỏ hơn n và với n đủ lớn:

$$\underbrace{\log_2\log_2\ldots\log_2}_{t+1 \text{ lần}}(2m_k) \geq \frac{n}{c\log_2 n} \geq 1.$$

Dẫn tới

$$\underbrace{\log_2\log_2\ldots\log_2}_{t+1 \text{ lần}}(\mathcal{P}(\mathcal{A}_t')) \geq \frac{n}{c\log_2 n}.$$

5. Biểu thức siêu chính quy $A_{t,n}$ được xây dựng từ biểu thức chính quy $\mathcal{A}_t'$:

$$\mathcal{A}_{t,n} = \mathcal{R}_t'\alpha 0^{t+2}\alpha\mathcal{R}_t'\alpha 0^{t+2}\alpha\mathcal{R}_t'(0 \vee 1 \vee \alpha)^\infty.$$

Biểu thức siêu chính quy này xác định ngôn ngữ siêu chính quy tương ứng, và như vậy độ phức tạp đoán nhận nó thỏa mãn

$$\underbrace{\log_2\log_2\ldots\log_2}_{t+1 \text{ lần}}(\mathcal{P}(\mathcal{A}_{t,n})) \geq \frac{n}{c\log_2 n}.$$

2.4. Định lý 2. *Với mỗi số tự nhiên tùy ý $n \geq 3$, có thể xây dựng được siêu ngôn ngữ chính quy trên bảng gồm 3 chữ cái, mà độ phức tạp otomat hữu hạn đoán nhận nó bằng 2^n.*

Chứng minh. Giả sử $\Sigma = \{a, b, c\}$. Ta xây dựng siêu nguồn G với đỉnh $\{v_1, v_2 \ldots v_n\}$ trong đó đỉnh vào là v_1 và tập đỉnh kết là $F = \{v_1\}$.

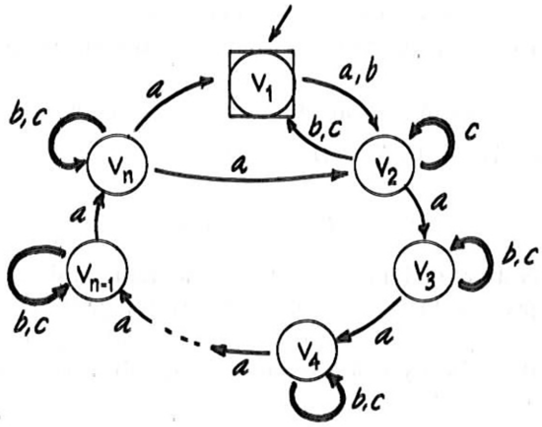
Gọi R là ngôn ngữ được sinh bởi G. Với mọi $\alpha \in \Sigma^*$, ký hiệu:

$S(\alpha) = \theta(v_1, \alpha)$ với $\alpha \neq \wedge$,

$S(\alpha) = \{v_1\}$ với $\alpha = \wedge$, hay $S(\wedge) = \{v_1\}$,

$E(\alpha) = \{\beta \in \Sigma^* \mid \beta \neq \wedge, \theta(v, \beta) \cap F \neq \emptyset$ với $v \in S(\alpha)\}$.

Nếu $S(\alpha) = \emptyset$ thì $E(\alpha) = \emptyset$.



Siêu nguồn G

Dễ thấy $\{\alpha\}E(\alpha)$ là tập từ (hữu hạn) được sinh ra từ nguồn G mà có đoạn đầu là α. Như vậy, ta thiết lập được sự tương ứng 1-1 giữa tập $E(\alpha)$ và $S(\alpha)$. Ta sẽ chứng minh các tập $S(\alpha)$ là phân biệt nhau từng đôi một và như vậy mỗi tập con từ 2^n tập con của tập các đỉnh của nguồn G được biểu diễn theo $S(\alpha)$.

Thật vậy, với mỗi tập con M không rỗng các đỉnh của nguồn G ta đặt $k(M) = (k_1, k_2, k_3)$ trong đó k_1 là số phần tử của M, k_2 là hiệu số giữa chỉ số lớn nhất và chỉ số nhỏ nhất của các đỉnh trong M, k_3 là chỉ số đỉnh nhỏ nhất trong M. Với $M = \{v_1\}$ thì $k(M) = (1, 0, 1)$. Ta thiết lập một quan hệ thứ tự giữa các tập M như sau: với 2 tập M với $k(M) = (k_1, k_2, k_3)$ và M' với $k(M') = (k'_1, k'_2, k'_3)$ trong đó $k_1 = k'_1, \ldots, k_i = k'_i, k_{i+1} < k'_{i+1}$ với i nào đó, $i \in \{0, 1, 2\}$ thì $k(M) < k(M')$.

Với tập con không rỗng tùy ý M các đỉnh của nguồn G mà đối với nó tất cả các tập con không rỗng M' của các đỉnh của G, thỏa mãn $k(M') < k(M)$, M' được biểu diễn bởi $S(\alpha)$ thì M cũng được biểu diễn bởi dạng $S(\alpha)$.

Trước hết ta dễ nhận thấy tất cả các tập con M chỉ có một đỉnh G (tức chỉ có một phần tử), đều được biểu diễn bởi dạng $S(\alpha)$. Chẳng hạn $S(aa) = \{v_3\}$.

Giả sử $M = \{\nu_{i_1}, \nu_{i_2}, \ldots, \nu_{i_p}\}$ trong đó $i_1 < i_2 < \ldots < i_p$.

Với $i_1 = 1$, $i_2 = 2$, ta xét tập $M' = \{\nu_{i_3-1}, \ldots, \nu_{i_p-1}, \nu_n\}$, nhận thấy $k'_1 < k_1$ nên $k(M') < k(M)$, do đó theo giả thiết thì $M' = S(\alpha)$ với α nào đó, dẫn tới $M = S(\alpha a)$ vì $\nu_{i_j-1} \xrightarrow{a} \nu_{i_j}$, và $\nu_n \xrightarrow{a} \nu_1$, $\nu_n \xrightarrow{a} \nu_2$.

Với $i_1 = 1$, $i_2 \geq 3$, ta xét tập $M' = \{\nu_1, \nu_{i_2-1}, \ldots, \nu_{i_p-1}\}$, nhận thấy $k'_2 < k_2$ nên $k(M') < k(M)$, do đó theo giả thiết thì $M' = S(\alpha)$ với α nào đó, dẫn tới $M = S(\alpha a)$ vì $\nu_{i_j-1} \xrightarrow{a} \nu_{i_j}$

Với $i_1 \geq 2$ ta xét tập $M' = \{\nu_{i_1-1}, \ldots, \nu_{i_p-1}\}$, nhận thấy $k'_3 < k_3$ nên $k(M') < k(M)$, do đó theo giả thiết thì $M' = S(\alpha)$ với α nào đó, dẫn tới $M = S(\alpha a)$ vì $\nu_{i_j-1} \xrightarrow{a} \nu_{i_j}$.

Ngoài ra do $S(\wedge) = \{v_1\}$ nên $S(\wedge c) = \emptyset$, tức là tập $\emptyset$ cũng biểu diễn được bởi dạng $S(\alpha)$. Vậy mọi tập con của tập các đỉnh của G đều được biểu diễn bởi dạng $S(\alpha)$.

Tiếp theo ta chứng minh các tập $E(\alpha)$ cũng khác nhau từng đôi một.

Giả sử $S(\alpha) = \{\nu_{i_1}, \nu_{i_2}, \ldots, \nu_{i_k}\}$ trong đó $i_1 < i_2 < \ldots < i_k$. Từ tập $E(\alpha)$ ta tách ra tập con $E'(\alpha)$ mà chỉ chứa các từ trên tập một chữ cái $\{a\}$, có độ dài từ 1 đến n. Dễ nhận thấy độ dài các từ trong $E'(\alpha)$ lần lượt bằng $n + 1 - i_1, n + 1 - i_2, \ldots, n + 1 - i_k$ với $i_k \neq n$ và bằng $n + 1 - i_1$, $n + 1 - i_2, \ldots, n + 1 - i_k$, n với $i_k = n$. Mặt khác $E(\alpha)$ còn chứa từ $baa...a$ ($n - 1$ chữ cái a) nếu $i_1 = 1$ và không chứa nó nếu $i_1 \neq 1$. Chứng tỏ các tập $E(\alpha)$ cũng khác nhau từng đôi một, tương ứng với 2^n tập $S(\alpha)$.

Nhận thấy $R \setminus \{\alpha\} = E(\alpha)$ nên theo Định lý Myhill-Nerode trên R, số trạng thái tối thiểu của otomat hữu hạn đơn định đoán nhận R không nhỏ hơn số lớp tương đương theo phép cắt đoạn đầu

$E(\alpha)$ là 2^n. Mặt khác số trạng thái tối thiểu của otomat hữu hạn đơn định tương đương với nguồn G không vượt quá 2^n. Dẫn tới số trạng thái tối thiểu của otomat hữu hạn đơn định đoán nhận R bằng 2^n.

Từ đó ta có thể xây dựng được siêu ngôn ngữ chính quy sinh ra bởi G như sau:

$$R = (((a \vee b)c^*(b \vee c))^*(((a \vee b)c^*(b \vee c))^*(a \vee b)c^*a(b \vee c)^* \ldots\ldots (b \vee c)a)^*)^\infty .$$

3. KẾT LUẬN

Vấn đề được đặt ra đã được giải quyết, điều đó có nghĩa là mối liên hệ giữa độ phức tạp đoán nhận siêu ngôn ngữ chính quy với độ sâu của phép đặt dấu lấy phần bù là quan hệ hàm mũ với giá trị nói chung là rất lớn. Liệu với những điều kiện nào thì độ phức tạp đoán nhận có thể giảm được trong khi độ sâu của phép đặt dấu lấy phần bù không thay đổi là vấn đề chúng tôi sẽ cố gắng giải quyết.

TÀI LIỆU THAM KHẢO

[1] Đặng Huy Ruận, Phùng Văn Ổn, Một số kết quả về lớp siêu ngôn ngữ, *Tuyển tập Hội nghị khoa học ngành Toán - Cơ - Tin học, trường Đại học Khoa học tự nhiên, Đại học Quốc gia Hà Nội*, tháng 4/1998.

[2] Đặng Huy Ruận, Phùng Văn Ổn, Độ phức tạp otomat hữu hạn đoán nhận siêu ngôn ngữ chính quy, *Hội thảo quốc gia về Tin học ứng dụng, Quy Nhơn*, tháng 8/1998.

[3] Dang Huy Ruan, On the complexity of a finite automaton corresponding to a generalized regular expression, *Dokl. Akad, Nauk SSSR*, T. 213, No. 1 (1973).

[4] Dang Huy Ruan, On the complexity of a finite automaton corresponding to special type of generating schemar, *Discrete Mathematics Banach Center Publications, Vol 7. PWN-Polish Scientific Publishers, Warzawa*, 1982.

[5] Đặng Huy Ruận - Độ phức tạp ô-tô-mát của các dãy biểu thức chính quy suy rộng, *Tạp chí Khoa học ĐHQG Hà Nội, KHTN* **XI** (1) (1995).

[6] V. B. Kyraxep, X. V. Alexen, A. X. Pokozin, *Introduction to Theory of Automata*, Nauka, 1988.

[7] D. B. Lupanov, Về việc so sánh hai loại nguồn hữu hạn, *Problem Cybernetic*, Vol. 9, Moskva, 1963.

Nhận bài ngày 10 - 6 - 1998
Nhận lại sau khi sửa ngày 26 - 2 - 1999

Đặng Huy Ruận - Khoa Toán - Cơ - Tin học,
Trường Đại học Khoa học Tự nhiên, ĐHQG Hà Nội.
Phùng Văn Ổn - Khoa Công nghệ thông tin, Trường Đại học Hàng hải.