Tạp chí Tin học và Điều khiển học, T. 14, S. 4 (1998) (25-30)

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

ĐẶNG HUY RUẬN(1), PHÙNG VĂN ỔN(2)

Abstract. We shall call the small possible number of states of a hyper-finite deterministic automaton, which recorgnizes the set of word L in the alphabet Σ, a complexity of a finite automaton of L, and denote by $\mathcal{P}(L)$. We have the following result:

For any hyper-generating schemar G, the complexity of a finite automaton of $L(G)$ is:

$$\mathcal{P}(L(G)) \leq 2^{2^{\cdot^{\cdot^{2^{|G|}}}}} \Big\} d(G) + 1 \text{ lần} \quad + 1$$

I. ĐẶT VẤN ĐỀ

Kết quả về lớp siêu ngôn ngữ chính quy cho ta thấy với mỗi siêu ngôn ngữ chính quy L, luôn tồn tại một siêu nguồn G mà $L(G) = L$, đồng thời tồn tại siêu otomat đơn định với trạng thái tối thiểu không quá $2^{|G|}$ đoán nhận $L(G)$ (gọi là độ phức tạp otomat hữu hạn đoán nhận G, ký hiệu là $\mathcal{P}(L)$) [1]. Mặt khác lớp siêu ngôn ngữ chính quy đóng với phép hợp, tích ghép, siêu lặp, giao và phép lấy phần bù. Điều đó có nghĩa là có thể nhận được một siêu ngôn ngữ chính quy bằng một số hữu hạn lần các phép tính: phép hợp, tích ghép, siêu lặp, giao và phép lấy phần bù trên một số hữu hạn các siêu ngôn ngữ chính quy, và nó được gọi là siêu ngôn ngữ chính quy suy rộng. Vấn đề đặt ra là với một siêu ngôn ngữ chính quy suy rộng tùy ý thì độ phức tạp đoán nhận nó sẽ như thế nào. Bài này nhằm giải quyết vấn đề đó.

II. SIÊU NGUỒN

Một đồ thị định hướng hữu hạn G trên bảng chữ cái Σ, có một đỉnh khởi đầu ν, một tập các đỉnh kết $\{\nu_1, \nu_2, ..., \nu_n\}$ và trên mỗi cạnh được ghi một từ rỗng (ký hiệu là ε, gọi là cung rỗng) hoặc một chữ cái a của bảng chữ cái Σ (gọi là cung cốt yếu) được gọi là *siêu nguồn* trên bảng chữ cái Σ. Đỉnh có ít nhất một cung cốt yếu đi tới gọi là đỉnh cốt yếu. Số đỉnh cốt yếu của G được kí hiệu là $|G|$ [4].

Siêu đường trong một siêu nguồn G là một dãy vô hạn $\pi : w_1, \rho_1, w_2, \rho_2, ...$, trong đó w_i, $i = 1, 2, 3, ...$ là các đỉnh của siêu nguồn G, còn ρ_i là cung của G đi từ đỉnh w_i đến đỉnh w_{i+1}. Siêu đường π lập nên *siêu từ* $[\pi] = a_{i_1}, a_{i_2} ...$ với a_{i_j} là chữ cái thuộc bảng chữ cái Σ, được ghi trên cung ρ_{i_j} trong đó $j = 1, 2, ...$ Đỉnh w_1 gọi là *điểm khởi đầu* của siêu đường π.

Đỉnh w được gọi là *đỉnh giới hạn* của siêu đường π nếu với tập vô hạn các chỉ số i thỏa mãn $w = w_i$. Tập các đỉnh giới hạn của siêu đường π được ký hiệu là $\lim(\pi)$. Hai siêu từ được gọi là trùng nhau nếu mọi đoạn đầu của chúng là như nhau và các siêu đường sinh ra chúng có cùng tập giới hạn.

Mỗi siêu nguồn G đều xác định một siêu ngôn ngữ (ký hiệu là $L(G)$) gồm tất cả các siêu từ $\alpha \in \Sigma^\infty$ để với siêu đường π nào đó với đỉnh khởi đầu ν (là đỉnh khởi đầu của G) thỏa mãn $[\pi] = \alpha$ và có ít nhất một đỉnh kết của siêu nguồn G thuộc tập $\lim(\pi)$.

Siêu nguồn G được gọi là *đơn định* nếu nó không chứa cung rỗng và trên hai cung khác nhau xuất phát từ cùng một đỉnh phải được ghi bởi hai chữ khác nhau.

Siêu nguồn G được gọi là *đầy đủ* nếu với bất kỳ chữ cái a nào thuộc Σ, từ mỗi đỉnh của G phải có ít nhất một cung xuất phát từ đó có ghi chữ cái a.

Siêu nguồn được gọi là *đơn định đầy đủ* nếu từ mỗi đỉnh của G có đúng n cung xuất phát từ đó và chúng được ghi các chữ cái khác nhau của bảng chữ cái Σ.

Cho hai siêu nguồn G_1 và G_2 không có đỉnh chung. Ta xét một số các phép tính sau:

1. Phép hợp hai siêu nguồn: Cho hai siêu nguồn G_1 và G_2. Ta xây dựng siêu nguồn G như sau: Tập đỉnh của G bao gồm tập đỉnh của G_1 và G_2. Ta thêm một đỉnh I_G, lấy làm đỉnh vào của G, từ I_G ta nối hai cung rỗng tới hai đỉnh vào của G_1 và G_2. Tập đỉnh kết của G bao gồm tất cả các đỉnh kết của G_1 và G_2. Siêu nguồn G được gọi là hợp của siêu nguồn G_1 và G_2 và ký hiệu là $G_1 \cup G_2$.

Bổ đề 1. *Hợp của G_1 và G_2 là một siêu nguồn G mà siêu ngôn ngữ sinh bởi G là hợp của các siêu ngôn ngữ sinh bởi G_1 và G_2 đồng thời số đỉnh cốt yếu của G bằng tổng số đỉnh cốt yếu của G_1 và G_2, tức là: $L(G) = L(G_1) \cup L(G_2)$ và $|G| = |G_1| + |G_2|$.*

2. Tích ghép của một nguồn với một siêu nguồn: Cho nguồn G_1 và siêu nguồn G_2. Ta xây dựng siêu nguồn G như sau: Tập đỉnh của G bao gồm tập đỉnh của G_1 và G_2. Đỉnh vào của G_1 được lấy làm đỉnh vào của G. Tập đỉnh kết của G_2 được lấy làm tập đỉnh kết của G. Từ tất cả các đỉnh kết của G_1 ta nối các cung rỗng tới đỉnh vào của G_2. Siêu nguồn G được gọi là tích ghép của nguồn G_1 và siêu nguồn G_2 (ký hiệu là $G_1.G_2$).

Bổ đề 2. *Tích ghép của nguồn G_1 và siêu nguồn G_2 là một siêu nguồn G mà siêu ngôn ngữ sinh bởi G tích ghép của ngôn ngữ từ hữu hạn sinh bởi G_1 và siêu ngôn ngữ sinh bởi siêu nguồn G_2 đồng thời số đỉnh cốt yếu của G bằng tổng số đỉnh cốt yếu của G_1 và G_2, tức là $L(G) = L(G_1).L(G_2)$ và $|G| = |G_1| + |G_2|$.*

3. Siêu lặp của một nguồn: Cho nguồn G_1. Ta xây dựng siêu nguồn G như sau: Tập đỉnh của G chính là tập đỉnh của G_1. Đỉnh vào của G_1 được lấy làm đỉnh vào của G và tập đỉnh kết của G_1 được lấy làm tập đỉnh kết của G. Từ tất cả các đỉnh kết của G ta nối các cung rỗng tới đỉnh vào của nó. Siêu nguồn G được gọi là siêu lặp của nguồn G_1 (ký hiệu là G_1^∞).

Bổ đề 3. *Siêu lặp của nguồn G_1 là một siêu nguồn G mà siêu ngôn ngữ sinh bởi G là siêu lặp của ngôn ngữ sinh bởi G_1, đồng thời số đỉnh cốt yếu của G bằng số đỉnh cốt yếu của G_1, tức là $L(G) = L(G_1)^\infty$ và $|G| = |G_1|$.*

Nhận xét: Nếu nguồn G là đầy đủ trên Σ và mọi số đỉnh của nó đều là đỉnh kết thì siêu lặp của G sẽ sinh ra Σ^∞.

4. Giao của hai siêu nguồn: Cho hai siêu nguồn G_1 và G_2. Ta xây dựng siêu nguồn G như sau: Tập đỉnh của G là tích Đề-các tập đỉnh của G_1 và G_2. Cặp đỉnh (I_{G_1}, I_{G_2}) được lấy làm đỉnh vào của G. Tích Đề-các tập đỉnh kết của G_1 và G_2 được lấy làm tập đỉnh kết của G. Các cung của G được xác định như sau:

Ta ký hiệu đỉnh đầu của cung t là $D(t)$, đỉnh cuối của cung t là $C(t)$.

a) Với mỗi bộ cung $t_1 \in G_1$ và $t_2 \in G_2$ mà trên đó đều ghi cùng chữ cái $a \in \Sigma$, ta xây dựng cung t tương ứng trên siêu nguồn G đi từ $D(t)$ đến $C(t)$ mà trên đó ghi chữ cái a, còn $D(t) = (D(t_1), D(t_2))$, $C(t) = (C(t_1), C(t_2))$.

b) Với mọi $D(t_i)$, $C(t_i)$ của siêu nguồn G_i $(i = 1, 2)$, trong G có cung rỗng từ $D(t) = ((D(t_1), D(t_2))$ sang $C(t) = (C(t_1), C(t_2))$ khi và chỉ khi trong G_i $(i = 1, 2)$, có cung rỗng đi từ $D(t_i)$ sang $C(t_i)$.

Siêu nguồn G được gọi là giao của G_1 và G_2 (ký hiệu là $G_1 \cap G_2$).

Bổ đề 4. *Giao của hai siêu nguồn G_1 và G_2 là một siêu nguồn G mà siêu ngôn ngữ sinh bởi G là giao của các siêu ngôn ngữ sinh bởi G_1 và G_2 đồng thời số đỉnh cốt yếu của G bằng tích các số đỉnh cốt yếu của G_1 và G_2, tức là $L(G) = L(G_1) \cap L(G_2)$ và $|G| = |G_1|.|G_2|$.*

5. Nguồn bù: Ta xây dựng siêu nguồn G là bù của siêu nguồn đơn định, đầy đủ G_1 (ký hiệu là $C(G_1)$ như sau: Trên siêu nguồn G_1, ta đổi tất cả các đỉnh kết thành không kết và ngược lại. Siêu nguồn nhận được chính là G.

Bổ đề 5. *Phép lấy phần bù của siêu nguồn đơn định, đầy đủ G_1 là một siêu nguồn G mà siêu ngôn ngữ sinh bởi G là phần bù của siêu ngôn ngữ sinh bởi G_1, đồng thời số đỉnh cốt yếu của G bằng số đỉnh cốt yếu của G_1, tức là $L(G) = CL(G_1)$ và $|G| = |G_1|$.*

III. SIÊU SƠ ĐỒ SINH

1. Siêu đồ thị sinh: Cho bảng chữ cái hữu hạn $\Sigma = \{a_1, a_2, ..., a_n\}$. Một đồ thị định hướng hữu hạn G với tập các đỉnh P_G, có một đỉnh vào I_G, một tập không rỗng các đỉnh kết $F_G = \{\nu_1, \nu_2, ..., \nu_k\}$ và trên mỗi cung t được ghi một tập các từ hữu hạn nào đó $M_G(t) \subseteq \Sigma^*$, được gọi là *siêu đồ thị sinh* trên bảng chữ cái Σ.

Dãy vô hạn $\pi = w_1, t_1, w_2, t_2, ...$ trong đó w_i, $i = 1, 2, ...$ là các đỉnh của G, t_i là cung đi từ đỉnh w_i đến đỉnh w_{i+1} được gọi là siêu đường trong siêu đồ thị sinh G. Siêu đường π lập nên siêu từ $[\pi] = \chi_{i_1} \chi_{i_2} ...$ với χ_{i_j} là các từ thuộc $M_G(t_{i_j})$ được ghi trên cung t_{i_j} trong đó $j = 1, 2, ...$ Đỉnh w_1 gọi là *điểm khởi đầu* của siêu đường π. Đỉnh w được gọi là *đỉnh giới hạn* của siêu đường π nếu với tập vô hạn các chỉ số i thỏa mãn $w = w_i$. Tập các đỉnh giới hạn của siêu đường π được ký hiệu là $\mathbf{lim}(\pi)$. Siêu từ $[\pi] = \chi_{i_1} \chi_{i_2} \ldots \in \Sigma^\infty$ gọi là *được sinh* bởi siêu đồ thị sinh G nếu $\lim(\pi) \cap F_G \neq \emptyset$ và điểm khởi đầu của siêu đường π chính là đỉnh vào I_G của đồ thị G. Tập hợp tất cả các siêu từ $[\pi]$ được sinh bởi G gọi là siêu ngôn ngữ được sinh ra bởi G và ký hiệu là $L(G)$ (ta ký hiệu $L^*(G)$ là tập các từ hữu hạn được sinh bởi đồ thị sinh G).

Nhận xét: Nếu mọi cung t của G mà $M_G(t) = \{\varepsilon\}$ hoặc $M_g(t) = \{a\}$, $a \in \Sigma$, thì siêu đồ thị sinh G là một siêu nguồn.

Giả sử G_1, G_2 là hai siêu đồ thị sinh không có đỉnh chung và t là một cung nào đó của G_1 nối từ đỉnh w_i tới w_j. Ta xây dựng siêu đồ thị sinh G từ G_1, G_2 bằng cách cắt bỏ cung t khỏi G_1; từ đỉnh w_i ta vẽ một cung rỗng tới đỉnh vào của G_2; từ mỗi đỉnh kết của G_2 ta vẽ các cung rỗng tới đỉnh w_j. Đỉnh vào của G_1 được nhận làm đỉnh vào của G, tập đỉnh kết của G_2 được nhận làm tập đỉnh kết của G. Siêu đồ thị sinh G như vậy được gọi là thế siêu đồ thị G_2 cho cung t của siêu đồ thị sinh G_1 và được ký hiệu là $G = [G_1]^t{}_{G_2}$.

Bổ đề 6. *Giả sử $G, G_1, G_2, ..., G_n$ là các siêu đồ thị sinh từng đôi một không có đỉnh chung; $t_1, t_2, ..., t_n$ là các cung của G mà với mỗi i, $i = 1, 2, ..., n$, có $M_G(t_i) = L^*(G_i)$. Ta xây dựng siêu đồ thị sinh G' bằng cách thế các cung t_i của G bởi các G_i. Khi đó ta có $L(G') = L(G)$ và $|G'| \leq |G| + \sum_{i=1}^{n} |G_i|$.*

2. Siêu sơ đồ sinh: Một siêu sơ đồ sinh G trên bảng chữ cái Σ là một dãy các siêu đồ thị sinh $G_1, G_2, ..., G_n$ trên Σ, ký hiệu là $G = (G_1, G_2, ..., G_n)$ thỏa mãn 3 điều kiện sau:

a) Các siêu đồ thị sinh G_1, G_2, ..., G_n từng đôi một không có đỉnh chung.

b) Với mỗi cung t của siêu đồ thị sinh G_i, $i = 1, 2, ...$ một trong các hệ thức sau được thỏa mãn:

1/ $M_{G_i}(t) = \{\varepsilon\}$, trong trường hợp này cung t được gọi là cung rỗng.

2/ $M_{G_i}(t) = \{a\}$, $a \in \Sigma$; trường hợp này cung t được gọi là cốt yếu.

3/ $M_{G_i}(t) = \bigcap_{j=1}^{s} L^*(G_{i_j})$, trường hợp này cung t được gọi là cung giao và phụ thuộc vào các siêu đồ thị $G_{i_1}, G_{i_2}, ..., G_{i_s}$, trong đó $1 \leq i_1 \leq i_2 \leq \cdots \leq i_s < i$.

4/ $M_{G_i}(t) = CL^*(G_j)$, $1 \leq j \leq i$; trường hợp này cung t được gọi là cung bù và phụ thuộc vào siêu đồ thị G_j.

c) Đối với mỗi siêu đồ thị sinh $G_1, G_2, ..., G_n$ luôn có một và chỉ một cung phụ thuộc nó.

Tập siêu đồ thị $L(G_n)$ được gọi là *siêu ngôn ngữ xác định bởi siêu sơ đồ sinh G* và ký hiệu là $L(G) : L(G) = L(G_n)$ [3].

Hai siêu sơ đồ sinh được gọi là *tương đương* nếu chúng cùng xác định một siêu ngôn ngữ.

Đỉnh w của siêu đồ thị sinh G_i được gọi là đỉnh cốt yếu nếu có ít nhất một cung cốt yếu đi tới nó. Số các đỉnh cốt yếu của siêu đồ thị sinh G_i được ký hiệu bởi $|G_i|$. Số tất cả các đỉnh cốt yếu của các siêu đồ thị sinh trong siêu sơ đồ sinh G, ký hiệu là $|G|$ và: $|G| = |G_1| + |G_2| + \cdots + |G_n|$.

Siêu đồ thị sinh G_i được gọi là phụ thuộc vào siêu đồ thị sinh G_j nếu chứa một cung nào đó phụ thuộc G_j hoặc nó phụ thuộc vào siêu đồ thị khác chứa cung phụ thuộc G_j.

Nếu siêu sơ đồ sinh G chỉ có một đồ thị sinh, tức $G = (G_1)$ thì G được gọi là siêu sơ đồ sinhđơn giản.

Nhận xét: Nếu G là một siêu sơ đồ sinh đơn giản thì trên mỗi cung t của nó, chỉ có thể chứa một chữ cái của Σ hoặc một từ rỗng, tức $M_G(t) = \{a\}$ hoặc $M_G(t) = \{\varepsilon\}$.

3. Độ sâu của phép đặt dấu lấy phần bù $d(G)$ của siêu sơ đồ sinh G, định nghĩa như sau:

a. Giả sử t là một cung của siêu đồ thị sinh G_i nào đó; ta định nghĩa đại lượng $d(t)$:
- Nếu t là cung rỗng hoặc cung cốt yếu thì $d(t) = 0$.
- Nếu t là cung bù phụ thuộc vào G_j và $d(G_j)$ đã xác định thì $d(t) = d(G_j) + 1$.
- Nếu t là cung giao phụ thuộc vào $G_{i_1}, G_{i_2}, \ldots, G_{i_s}$ trong đó $1 \le i_1 \le i_2 \le \ldots \le i_s < i$, thì $d(t) = \max\big(d(G_{i_1}), d(G_{i_2}), \ldots, d(G_{i_s})\big)$.

Khi đó $d(G_i) = \max(d(t)), t \in G_i$.

b. $d(G) = \max d(G_i), i = 1, 2, \ldots, n$.

Định lý 1. *Với mỗi siêu sơ đồ sinh tùy ý G, ta luôn xây dựng được một siêu sơ đồ sinh đơn giản G' tương đương với nó sao cho:*

$$|G'| \le 2^{2^{\cdot^{\cdot^{\cdot^{2^{|G|}}}}}} \quad d(G)+1 \text{ lần}$$

Chứng minh: Ta chứng minh quy nạp theo số các siêu đồ thị sinh trong G.

Để thuận tiện, ta ký hiệu:

$$h(d(G), |G|) = 2^{2^{\cdot^{\cdot^{\cdot^{2^{|G|}}}}}} \quad d(G)+1 \text{ lần}$$

• Giả sử $n = 1$, khi đó G là siêu sơ đồ sinh đơn giản, tức G là một siêu nguồn. G' nhận được bằng cách đơn định hóa G, và như vậy $|G'| = |G| \le 2^{|G|}$, đồng thời $L(G') = L(G)$.

• Giả sử điều khẳng định đúng với mọi siêu sơ đồ sinh có số siêu đồ thị sinh nhỏ hơn hoặc bằng $n-1$, ta chứng minh điều khẳng định đúng với n.

Giả sử $G = (G_1, G_2, \ldots, G_n)$, $n \ge 2$ và giả sử trong G_n có p cung giao: $g_1, g_2, \ldots, g_p$ và q cung bù $b_1, b_2, \ldots, b_q$ $(p, q < n)$.

1. Với mỗi i $(1 \le i \le p)$, g_i là cung giao phụ thuộc vào các siêu đồ thị $G_{i_1}, G_{i_2}, \ldots, G_{i_m}$ $(i_{m_i} \ge 2)$ và $M_{G_n}(g_i) = \bigcap\limits_{j=1}^{m_i} L^*(G_{i_j})$.

Đối với mỗi j $(1 \le j \le m_i)$:

- Ta lấy các đồ thị sinh trong G mà G_{i_j} phụ thuộc, kể cả chính nó. Lập siêu sơ đồ sinh tương ứng G_{1,i_j}. Như vậy $L(G_{1,i_j}) = L(G_{i_j})$.

- Do số siêu đồ thị sinh trong siêu sơ đồ sinh G_{1,i_j} nhỏ hơn hoặc bằng p nên thỏa mãn giả thiết quy nạp, và ta có thể xây dựng được một siêu sơ đồ sinh đơn giản G'_{1,i_j} tương đương: $L(G'_{1,i_j}) = L(G_{1,i_j}) = L(G_{i_j})$ và:

$$|G'_{1,i_j}| \le h(d(G_{1,i_j}), G'_{1,i_j}) \quad (1 \le i \le p \text{ và } 1 \le j \le m_i).$$

Từ các sơ đồ sinh đơn giản vừa xây dựng $G'_{1,i_1}, G'_{1,i_2}, ..., G'_{1,i_m}$ (là những siêu nguồn), ta xây dựng siêu sơ đồ sinh đơn giản $W_{1,i}$ là giao của chúng và theo bổ đề 4 ta có:

$$L(W_{1,i}) = \bigcap_{j=1}^{m_i} L^*(G'_{1,i_j}) = \bigcap_{j=1}^{m_i} L^*(G_{1,i_j}) = M_{G_n}(g_i)$$

và

$$|W_{1,i}| \le \prod_{j=1}^{m_i} G'_{1,i_j} \le \prod_{j=1}^{m_i} h(d(G_{1,i_j}), |G_{1,i_j}|) \le \prod_{j=1}^{m_i} h(d(G), |G_{1,i_j}|). \tag{1}$$

2. Với mỗi i $(1 \le i \le q)$, b_i là cung bù phụ thuộc vào siêu đồ thị G_i và $M_{G_n}(b_i) = CL(G_i)$.

Lấy các siêu đồ thị sinh trong G mà G_i phụ thuộc, kể cả chính nó. Lập siêu sơ đồ sinh tương ứng $G_{2,i} = (G_{i_1}, G_{i_2}, ..., G_{i_{m_i}})$, $n \ge m_i \ge 1$. Như vậy $L(G_{2,i}) = L(G_i)$. Do số siêu đồ thị trong siêu sơ đồ $G_{2,i}$ nhỏ hơn hoặc bằng $(n-1)$ nên thỏa mãn giả thiết quy nạp, và ta có thể xây dựng được một siêu sơ đồ sinh đơn giản G'_{2i} tương đương: $L(G'_{2,i}) = L(G_{2,i}) = L(G_i)$ và:

$$|G'_{2,i}| \le h(d(G_{2,i}), |G_{2,i}|).$$

Với siêu sơ đồ sinh đơn giản $G'_{2,i}$, theo bổ đề 5, ta luôn xây dựng được siêu sơ đồ sinh đơn giản $W_{2,i}$ sao cho:

$$L(W_{2,i}) = CL(G'_{2,i}) = CL(G_{2,i}) = M_{G_n}(b_i) \text{ với } |W_{2,i}| \le 2^{|G'_{2,i}|}.$$

Kết hợp với kết quả trên, ta có: $|W_{2,i}| \le h(d(G_{2,i}) + 1, |G_{2,i}|)$.

Vì b_i là cung bù phụ thuộc G_i nên $d(b_i) = d(G_i) + 1 = d(G_{2,i}) + 1$, dẫn tới $d(G_{2,i}) = d(b_i) - 1 \le d(G) - 1$. Thay vào biểu thức trên ta được:

$$|W_{2,i}| \le h(d(G), |G_{2,i}|). \tag{2}$$

3. Thế các siêu sơ đồ sinh đơn giản $W_{1,i}$, $W_{2,j}$ với $i = 1, ..., p$, $j = 1, ..., q$ vào siêu đồ thị sinh G_n thay cho các cung tương ứng g_i, b_j. Khi đó ta nhận được siêu đồ thị sinh đơn giản G' mà $L(G') = L(G_n) = L(G)$. Theo bổ đề 6 ta có:

$$|G'| = |G_n| + \sum_{i=1}^{p} |W_{1,i}| + \sum_{i=1}^{q} |W_{2,i}|.$$

Thế các kết quả (1) và (2) vào ta được:

$$|G'| \le |G_n| + \sum_{i=1}^{p} \prod_{j=1}^{m_i} h(d(G), |G_{1,i_j}|) + \sum_{i=1}^{q} h(d(G), |G_{2,i}|). \tag{3}$$

Không mất tính tổng quát, ta có thể giả thiết các siêu sơ đồ sinh $G'_{1,i}$, $G'_{2,i}$ đều có ít nhất một đỉnh cốt yếu (vì trong trường hợp ngược lại các tập từ tương ứng của chúng $M_{G_n}(g_{i_j})$, $M_{G_n}(b_i)$ sẽ là tập rỗng hoặc chỉ gồm một từ trống, khi đó ta có thể đơn giản siêu sơ đồ sinh G bằng cách bỏ đi các cung tương ứng hoặc thay chúng bằng các cung rỗng). Khi đó ta có:

$$|G_{1,i_j}| \ge 1 \ (i = 1, 2, ..., p,\ j = 1, 2, ..., m_i); \quad |G'_{2,i}| \ge 1 \ (i = 1, 2, ..., q). \tag{4}$$

- Trường hợp trong G có ít nhất một cung bù, khi đó $d(G) \ge 1$ nên từ (4) ta có:

$$\sum_{i=1}^{p} \prod_{j=1}^{m_i} h(d(G), |G_{1,i_j}|) \le h\Big(d(G), \sum_{i=1}^{p} \prod_{j=1}^{m_i} |G_{1,i_j}|\Big).$$

Tương tự ta có:
$$\sum_{i=1}^{q} h(d(G), |G_{2,i}|) \le h\Big(d(G), \sum_{i=1}^{q} |G_{2,i}|\Big).$$

Từ (3) và các kết quả trên ta có:

$$|G'| \le |G_n| + h\Big(d(G), \sum_{i=1}^{p}\prod_{j=1}^{m_i}|G_{1,i_j}|\Big) + h\Big(d(G), \sum_{i=1}^{q}|G_{2,i}|\Big)$$
$$\le h\Big(d(G), |G_n| + \sum_{i=1}^{p}\prod_{j=1}^{m_i}|G_{1,i_j}| + \sum_{i=1}^{q}|G_{2,i}|\Big) \le h(d(G), |G|)$$

- Trường hợp trong G không có cung bù, khi đó $d(G) = 0$; ta có:

$$|G'| \le |G_n| + h\Big(d(G), \sum_{i=1}^{p}\prod_{j=1}^{m_i}|G_{1,i_j}|\Big) \le h\Big(d(G), |G_n| + \sum_{i=1}^{p}\prod_{j=1}^{m_i}|G_{1,i_j}|\Big) \le 2^{|G|}.$$

Định lý được chứng minh.

IV. ĐỘ PHỨC TẠP OTOMAT HỮU HẠN ĐOÁN NHẬN SIÊU NGÔN NGỮ CHÍNH QUY

Định lý 2. *Với mỗi siêu đồ sinh G, độ phức tạp otomat hữu hạn đoán nhận siêu ngôn ngữ chính quy sinh bởi G thỏa mãn:*

$$P(L(G)) \le h(d(G)+1, |G|) + 1.$$

Chứng minh: Với mỗi siêu sơ đồ sinh G. theo định lý 1 ta luôn xây dựng được một siêu sơ đồ sinh đơn giản G' thỏa mãn $L(G) = L(G')$ và $|G'| \le h(d(G), |G|)$. Đơn định hóa G' ta nhận được siêu nguồn G'' tương đương mà $L(G'') = L(G')$ và

$$|G''| \le 2^{|G'|} + 1 = h(d(G)+1, |G|) + 1.$$

Do G'' đơn định nên tồn tại siêu otomat hữu hạn đơn định, có số trạng thái bằng $|G''|$, tương đương với G'', và do đó nó đoán nhận $L(G'')$, tức đoán nhận $L(G)$. Ta có điều phải chứng minh.

V. KẾT LUẬN

Các kết quả trên cho ta đánh giá cận trên của độ phức tạp đoán nhận siêu ngôn ngữ chính quy. Nó cho một khả năng hữu hạn hóa khi làm việc với những ngôn ngữ từ vô hạn. Tuy nhiên giá trị cận trên là rất lớn, chúng tôi sẽ tiến hành xem xét những trường hợp riêng với những hạn chế nào đó để có thể có những kết quả tốt hơn.

TÀI LIỆU THAM KHẢO

[1] Đặng Huy Ruận, Phùng Văn Ổn, Một số kết quả về lớp siêu ngôn ngữ chính quy, *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] Dang Huy Ruan, On the complexity of a finite automaton corresponding to a gereralized regular expression, *Dokl. Akad. Nauk SSSR*, Tom 213, No. 1 (1973).

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

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

Nhận bài ngày 5 - 2 - 1998

(1) *Khoa Toán - Cơ - Tin học,*
Trường Đại học Khoa học tự nhiên – ĐHQG Hà Nội.

(2) *Khoa học Công nghệ thông tin, Trường Đại học Hàng hải.*