Tạp chí Tin học và Điều khiển học, T.26, S.4 (2010), 301 - 311

MỞ RỘNG MÃ VÀ THUẬT TOÁN KIỂM ĐỊNH MÃ LUÂN PHIÊN VÀ MÃ CỦA CÁC TỪ ĐỊNH BIÊN

HỒ NGỌC VINH[1] , PHAN TRUNG HUY[2] , ĐỖ LONG VÂN[3]

[1]*Khoa Công nghệ thông tin, Trường Đại học Sư phạm Kỹ thuật Vinh*
[2]*Khoa Toán - Tin ứng dụng, Trường Đại học Bách khoa Hà nội*
[3]*Viện Toán học, Viện Khoa học và Công nghệ Việt Nam*

Abstract. In this paper, we consider properties of +-unambiguous products, alternative codes and languages of bounded words ($\diamond$-languages). A new type of codes ($\diamond$-codes or codes of bounded words) is defined and an algorithm to determine whether a $\diamond$-recognizable language is a $\diamond$-code is established. This provides a new algorithm to verify for alternative codes whose complexity is smaller than the complexity of the previous one. A relationship between codes, alternative codes and $\diamond$-codes is considered. This shows that $\diamond$-codes is an extension of traditional codes and alternative codes.

Tóm tắt. Bài báo xem xét một số tính chất của +-tích không nhập nhằng, mã luân phiên và ngôn ngữ của các từ định biên ($\diamond$-ngôn ngữ). Từ đó, xây dựng một hệ mã mới - mã của các từ định biên ($\diamond$-mã) cùng với thuật toán kiểm định một $\diamond$-ngôn ngữ đoán nhận được có là $\diamond$-mã hay không. Nhờ thuật toán kiểm định $\diamond$-mã, cho phép nhận được một thuật toán kiểm định mã luân phiên mới, làm giảm rõ rệt về độ phức tạp. Ngoài ra, bài báo còn đề xuất một sơ đồ phân lớp giữa ba lớp mã - mã, mã luân phiên và $\diamond$-mã, cho thấy $\diamond$-mã được xem như là sự mở rộng của mã và mã luân phiên.

1. MỞ ĐẦU

Tích không nhập nhằng của các từ hữu hạn cũng như của các từ vô hạn đã mang lại nhiều tính chất lý thú và đang được nhiều người quan tâm nghiên cứu bởi chúng có quan hệ gần gũi với mã. Để làm giàu thêm các tính chất của lý thuyết mã, gần đây đã có nhiều kết quả khi nghiên cứu tích không nhập nhằng trong quan hệ với mã, otomat, đại số ... Khái niệm tích không nhập nhằng được đề xuất bởi M. P. Schützenberger [1] và được mở rộng bởi J. E. Pin [2, 3]. Trong [4], Pascal Weil đã đưa vào sử dụng otomat không nhập nhằng, vị nhóm quan hệ không nhập nhằng để làm công cụ thiết lập tích quan hệ $X = Y \circ Z$ với Y, Z là mã hữu hạn trong quan hệ với lý thuyết đa tạp của ngôn ngữ. Trong [5], P. T. Huy và Đ. L. Vân đã thiết lập được kết quả biểu diễn ω-ngôn ngữ chính quy của các từ hữu hạn được đoán nhận bởi $\mathcal{V}$-otomat Büchi không nhập nhằng.

Trong bài báo này, chúng tôi nhắc lại khái niệm +-tích không nhập nhằng, tích luân phiên, mã luân phiên, mã luân phiên chẵn của hai ngôn ngữ X, Y thuộc A^* được đề xuất bởi P. T. Huy và V. T. Nam [6]; một số tính chất của +-tích không nhập nhằng, mã luân

phiên, mã luân phiên chẵn, đặc trưng cần và đủ để một cặp (X, Y) là mã luân phiên được xem xét bởi H. N. Vinh, V. T. Nam và P. T. Huy [7]. Từ đó, đưa ra sơ đồ quan hệ với mã truyền thống (Nhận xét 2.1), cho thấy mã luân phiên là sự mở rộng của mã và là lớp con của mã luân phiên chẵn. Trong sự phân lớp này, lớp mã là lớp con nhỏ nhất, lớp mã luân phiên là lớp nằm giữa và lớp lớn nhất là lớp mã luân phiên chẵn.

Theo [8, 9], một dạng ngôn ngữ mới, ngôn ngữ của các từ định biên ($\diamond$-ngôn ngữ) được giới thiệu và từ đó cho phép thiết lập kết quả cơ bản (Định lý 3.1) về sự tương đương giữa tính đoán nhận được của $\diamond$-ngôn ngữ theo $\diamond$-otomat và theo $\diamond$-đồng cấu vị nhóm. Một hình thức mã mới của các từ định biên ($\diamond$-mã), thuật toán kiểm định $\diamond$-mã khi nó là $\diamond$-ngôn ngữ đoán nhận được bởi $\diamond$-otomat hữu hạn được thiết lập. Từ đó, cho phép nhận được một kết quả cơ sở của bài báo về sự phân lớp (Định lý 4.1) giữa ba loại mã – mã (truyền thống), mã luân phiên và $\diamond$-mã, cho thấy $\diamond$-mã được xem như là một sự mở rộng của mã và mã luân phiên. Cũng trong bài báo này, thuật toán kiểm định mã luân phiên mới thông qua thuật toán kiểm định $\diamond$-mã được đề xuất cùng với sự đánh giá độ phức tạp tính toán (Hệ quả 4.2) để làm cơ sở so sánh với thuật toán kiểm định mã luân phiên trong [7]. Tiếp cận mới này cho phép giảm rõ rệt độ phức tạp so với thuật toán cũ (dựa vào Định lý 4.3 làm cơ sở toán học của thuật toán).

Trước hết, ta nhắc lại các ký hiệu và khái niệm được trình bày trong [10, 11]. Cho A là bảng chữ cái hữu hạn. Ký hiệu A^* là vị nhóm tự do sinh bởi A, với phép toán tích ghép, phần tử đơn vị là ε (từ rỗng) và $A^+ = A^* - \{\varepsilon\}$. Giả sử $u, v \in A^*$, ta nói u là *một khúc* (*khúc đầu, khúc đuôi*) của v nếu tồn tại $x, y \in A^*$ sao cho $v = xuy$ (tương ứng $v = uy, v = xu$). Một khúc (khúc đầu, khúc đuôi) là thực sự nếu $xy \neq \varepsilon$ (tương ứng $y \neq \varepsilon, x \neq \varepsilon$). Số tất cả các xuất hiện của các chữ trong từ u là độ dài của u, ký hiệu là $|u|$, quy ước $|\varepsilon| = 0$. Cho $X \subseteq A^+$, một phân tích của từ $w \in A^*$ trong X là một dãy $\{w_1, w_2, ..., w_n\}$, với $n \geq 1$ sao cho $w = w_1 w_2 \dots w_n, w_i \in X$; X được gọi là *mã* nếu mọi từ $w \in A^+$ có nhiều nhất một cách phân tích thành các từ trong X. Giả sử $X, Y \subseteq A^*$, ta gọi *thương trái* (*thương phải*) của X với Y là ngôn ngữ $Y^{-1}X$ (tương ứng XY^{-1}) được xác định bởi $Y^{-1}X = \{w \in A^* \mid yw \in X, y \in Y\}$ và $XY^{-1} = \{w \in A^* \mid wy \in X, y \in Y\}$.

2. MÃ LUÂN PHIÊN VÀ THUẬT TOÁN KIỂM ĐỊNH MÃ LUÂN PHIÊN

Phần này nhắc lại một số khái niệm và kết quả theo [6, 7]. Cho bảng chữ A và $X, Y \subseteq A^+$. Khi đó, tích XY được gọi là *tích không nhập nhằng* (hay (X, Y) *là tích không nhập nhằng*) nếu:

$$x, x' \in X,\ y, y' \in Y :\ xy = x'y' \ \Rightarrow\ x = x' \text{ và } y = y'.$$

Lưu ý rằng, mã cũng có thể được xem như một trường hợp đặc biệt của tích không nhập nhằng (X là mã thì (X, X) là tích không nhập nhằng). Chúng ta mở rộng khái niệm tích không nhập nhằng qua nhiều lần lặp tích, gọi là *+-tích không nhập nhằng*.

Cho $X, Y \subseteq A^+$, ta gọi *+-tích* của X, Y là tập được định nghĩa bởi

$$(XY)^+ = (XY) \cup (XY)^2 \cup (XY)^3 \dots$$

Ta định nghĩa *sự phân tích luân phiên* theo cặp (X, Y) của hai tập ngôn ngữ $X, Y \subseteq A^+$ như sau:

Định nghĩa 2.1. Cho bảng chữ A và $X, Y \subseteq A^+, w \in A^+$. Khi đó ta nói rằng:

(i) Từ w thừa nhận một phân tích luân phiên theo (X, Y) nếu $w = u_1u_2 ... u_n$ $(n \geq 2)$, trong đó, $u_1 \in X$, nếu $u_i \in X$ thì $u_{i+1} \in Y$ và nếu $u_i \in Y$ thì $u_{i+1} \in X$, với $\forall i = 1, .., n-1$.

(ii) Từ w thừa nhận một phân tích luân phiên chẵn theo (X, Y) nếu $w = u_1u_2 ... u_n$ $(n \geq 2)$, trong đó, $u_1 \in X, u_n \in Y$, nếu $u_i \in X$ thì $u_{i+1} \in Y$ và nếu $u_i \in Y$ thì $u_{i+1} \in X$, với $\forall i = 1, ..., n-1$ và n chẵn.

(iii) Từ w thừa nhận một phân tích luân phiên theo X, Y nếu w thừa nhận một phân tích luân phiên theo (X, Y) hoặc (Y, X).

Với $X, Y \subseteq A^+$, ký hiệu $S_{X,Y}$ là tập tất cả các từ $w \in A^+$ thừa nhận ít nhất một tích luân phiên theo X, Y. Khi đó $S_{X,Y} \subseteq (X \cup Y)^+$.

Ví dụ 2.1. Cho $X = \{a, ba\}$ và $Y = \{b, aba\}$. Khi đó, từ $w = abababaaba$ có hai phân tích luân phiên theo X, Y như sau:

f_1 : $(a).(b).(a).(b).(a).(aba)$ là một phân tích luân phiên chẵn theo (X, Y) và

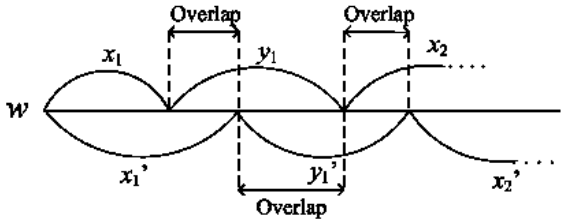
f_2 : $(aba).(ba).(aba)$ là một phân tích luân phiên theo (Y, X).

Định nghĩa 2.2. Cho $X, Y \subseteq A^+$. Khi đó, +-tích $(XY)^+$ được gọi là có +-tích không nhập nhằng (hay (X, Y) là +-tích không nhập nhằng) khi và chỉ khi: $\forall m, n \geq 2, x_1, x_2, ..., x_n$, $x'_1, x'_2, ..., x'_m \in X$, $y_1, y_2, ..., y_n$, $y'_1, y'_2, ..., y'_m \in Y$,

$$x_1y_1x_2y_2 ... x_ny_n = x'_1y'_1x'_2y'_2 ... x'_my'_m \Rightarrow m = n, x_i = x'_i, y_i = y'_i, \forall i = 1, ..., n.$$

Nói cách khác, sự phân tích của một từ w luân phiên trong X, Y dưới dạng $x_1y_1x_2y_2 ... x_ny_n$ nếu có thì sẽ là duy nhất.

Chú ý rằng, nếu (X, Y) là +-tích không nhập nhằng thì (X, Y) là tích không nhập nhằng. Đặc biệt, trong trường hợp cặp (X, Y) là *+-tích nhập nhằng* thì tồn tại một từ $w \in A^+$ sao cho các Overlap của hai phân tích của từ w là khác ε. (xem Hình 2.1).



Hình 2.1. Các bao trùm của hai phân tích của từ w

Dựa trên khái niệm +-tích không nhập nhằng và khái niệm phân tích luân phiên ở trên, cho phép ta định nghĩa một lớp mã mới như sau:

Định nghĩa 2.3. Cho $X, Y \subseteq A^+$. Khi đó:

(i) Cặp (X, Y) được gọi là mã luân phiên nếu mỗi từ $w \in A^+$ thừa nhận không quá một tích luân phiên theo X, Y.

(*ii*) Cặp (X, Y) được gọi là mã luân phiên chẵn nếu mỗi từ $w \in A^+$ thừa nhận không quá một tích luân phiên chẵn theo (X, Y).

Ví dụ 2.2. Cho $X = \{a, aa\}$, $Y = \{ab, b\}$, dễ thấy $XY = \{aab, ab, aaab\}$ là mã prefix nhưng cặp (X, Y) không là mã luân phiên, với từ $w = ababaab \in A^+$ thừa nhận hai sự phân tích luân phiên khác nhau theo X, Y là:

$$w = (a).(b).(a).(b).(aa).(b) = (ab).(a).(b).(aa).(b)$$

Mối liên hệ giữa +-tích không nhập nhằng, tích luân phiên và mã luân phiên được thể hiện qua định lý sau:

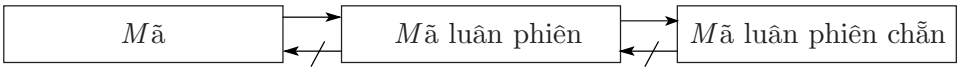
Định lý 2.1. *Cho $X, Y \subseteq A^+$. Khi đó:*

(*i*) *Cặp (X, Y) là mã luân phiên chẵn khi và chỉ khi cặp (X, Y) là +-tích không nhập nhằng.*

(*ii*) *Cặp (X, Y) là +-tích không nhập nhằng khi và chỉ khi $Z = XY$ là mã và (X, Y) là tích không nhập nhằng.*

(*iii*) *Nếu cặp (X, Y) là mã luân phiên thì cặp (X, Y) là +-tích không nhập nhằng, ngược lại không đúng.*

Nhận xét 2.1. Từ Định nghĩa 2.3 và Định lý 2.1, dễ thấy rằng: Mã X là một trường hợp đặc biệt của mã luân phiên (X, Y), với $Y = X$ và mã luân phiên là trường hợp đặc biệt của mã luân phiên chẵn. Nhưng ngược lại không đúng (xem Hình 2.2).



Hình 2.2. Quan hệ giữa các lớp mã

Bài toán cơ bản khi nghiên cứu về lớp mã là bài toán kiểm tra tính chất mã của ngôn ngữ (chính quy) cho trước. Kết quả sau thể hiện đặc trưng tổ hợp của mã luân phiên, đặc trưng cần và đủ đối với một cặp (X, Y) cho trước có là mã luân phiên hay không. Từ đó cho phép thiết lập thuật toán kiểm định mã luân phiên khi X, Y là các ngôn ngữ chính quy.

Định lý 2.2. *Cho $X, Y \subseteq A^+$. Khi đó, (X, Y) là mã luân phiên khi và chỉ khi bốn điều kiện sau đồng thời được thỏa mãn:*

(*i*) XY *là mã và* $X^{-1}X \cap YY^{-1} - \{\varepsilon\} = \emptyset$;

(*ii*) $Y^{-1}(XY)^+ \cap (XY)^+ = \emptyset$;

(*iii*) $(XY)^+X^{-1} \cap (XY)^+ = \emptyset$;

(*iv*) $(XY)^+ \cap (YX)^+ = \emptyset$.

3. ◇-NGÔN NGỮ ĐOÁN NHẬN ĐƯỢC

Phần này nhắc lại một số khái niệm liên quan đến từ định biên và ◇-ngôn ngữ trong [8] và một số kết quả cơ sở trong [9]. Cho A là bảng chữ cái hữu hạn và $B = \{0, 1\}$ là tập biên. Ta xây dựng các tập gồm các phần mở rộng của các từ trong A^* được xác định bởi:

$$\begin{aligned}
A_\diamond &= \{(i,a,j) \mid a \in A \text{ hoặc } a=\varepsilon,\ i,j \in B\}\\
A^*_\diamond &= \{(i,w,j) \mid w \in A^*,\ i,j \in B\} \cup \{\theta, e\}\\
A^*_\lhd &= \{(0,w,1) \mid w \in A^*\} \cup \{\theta, e\};\\
A^*_\rhd &= \{(1,w,0) \mid w \in A^*\} \cup \{\theta, e\}\\
A^*_\triangle &= \{(0,w,0) \mid w \in A^*\} \cup \{\theta, e\};\\
A^*_\triangledown &= \{(1,w,1) \mid w \in A^*\} \cup \{\theta, e\}
\end{aligned}$$

Mỗi bộ $(i,w,j), w \in A^*$ được gọi là một $\diamond$-*từ* (*từ định biên với biên i, j*), là mở rộng của từ w. Trong đó $e, \theta \notin A^*$ là hai phần tử mới đóng vai trò như là phần tử zero và phần tử đơn vị của tập các $\diamond$-từ trên $A^*_\diamond$ khi ta trang bị một phép toán (gọi là *tích biên*) trên $A^*_\diamond$ như sau: với $\forall x_1 = (i_1, w_1, j_1)$, $x_2 = (i_2, w_2, j_2) \in A^*_\diamond$,

$$x_1.x_2 = \begin{cases} (i_1, w_1w_2, j_2) & j_1 = i_2 \\ \theta & j_1 \neq i_2 \end{cases}$$

và $\forall x \in A^*_\diamond,\ x.\theta = \theta.x = \theta,\ x.e = e.x = x$

Dễ thấy $A^*_\diamond$ lập thành một vị nhóm với phép toán tích biên, có phần tử zero là θ và đơn vị là e. Ta gọi $A^*_\diamond$ là $\diamond$-*vị nhóm* (*kết hợp với bảng chữ* A). Ta quy ước rằng, các ký hiệu $\diamond, \lhd, \rhd, \triangle, \triangledown$ như là các toán tử mở rộng của A^* thành $A^*_\diamond$, $A^*_\lhd$, $A^*_\rhd$, $A^*_\triangle$, $A^*_\triangledown$ tương ứng.

Tiếp theo, ta định nghĩa các tập:

$$\begin{aligned}
\varepsilon_\diamond &= \{(i,\varepsilon,j) \mid i,j \in B\}; & A^*_{\diamond\setminus\varepsilon} &= A^*_\diamond - \varepsilon_\diamond\\
A^+_\diamond &= A^*_\diamond - \{e\}; & A^+_{\diamond\setminus\varepsilon} &= A^+_\diamond - \varepsilon_\diamond\\
A^+_\lhd &= A^*_\lhd - \{e\}; & A^+_{\lhd\setminus\varepsilon} &= A^+_\lhd - \varepsilon_\diamond\\
A^+_\rhd &= A^*_\rhd - \{e\}; & A^+_{\rhd\setminus\varepsilon} &= A^+_\rhd - \varepsilon_\diamond\\
A^+_\triangle &= A^*_\triangle - \{e\}; & A^+_{\triangle\setminus\varepsilon} &= A^+_\triangle - \varepsilon_\diamond\\
A^+_\triangledown &= A^*_\triangledown - \{e\}; & A^+_{\triangledown\setminus\varepsilon} &= A^+_\triangledown - \varepsilon_\diamond
\end{aligned}$$

Một tập $L \subseteq A^*_\diamond$ được gọi là một *ngôn ngữ mở rộng* ($\diamond$-*ngôn ngữ*) trên A. Với $x = (i,w,j) \in A^*_\diamond$, nếu không sợ hiểu nhầm thì ta cũng sử dụng ký hiệu $|x|$ là độ dài của $\diamond$-từ x, theo nghĩa $|x| = 0$ nếu $x \in \varepsilon_\diamond$ và $|x| = |w|$ nếu khác, ngoài ra $|\theta| = +\infty, |e| = 0$.

Giả sử $X, Y \subseteq A^*_\diamond$, ta gọi *thương trái* (*thương phải*) của X với Y là $Y^{-1}X$ (tương ứng XY^{-1}) trên $A^*_\diamond$ được xác định bởi:

$$Y^{-1}X = \{x \in A^*_\diamond \mid \exists y \in Y : y.x \in X\} \text{ và } XY^{-1} = \{x \in A^*_\diamond \mid \exists y \in Y : x.y \in X\}$$

Ta định nghĩa *phép chiếu* $Proj : A^*_\diamond \to A^* \cup \{0\}, 0 \notin A^*$, là hàm được xác định bởi :

$$Proj(e) = \varepsilon,\ Proj(\theta) = 0 \quad \text{và} \quad Proj(i,w,j) = w.$$

Cho $X \subseteq A^*_\diamond$ và $x \in A^*_\diamond$. Ta nói rằng x *thừa nhận một phân tích* (với tích biên) trong X nếu x có thể biểu diễn dưới dạng $x = x_1.x_2 \ldots x_n$, với $n \geq 1, \forall x_i \in X, \forall i = 1, \ldots, n$

Ví dụ 3.1. Cho $X = \{(1, a, 0), (0, ab, 1), (1, ba, 0)\}$. Khi đó $x = (1, baabaabba, 0)$ thừa nhận một sự phân tích trong X là:

$$(1, baabaabba, 0) = (1, ba, 0).(0, ab, 1).(1, a, 0).(0, ab, 1).(1, ba, 0)$$

Định nghĩa 3.1. Cho M là một vị nhóm bất kỳ có chứa phần tử đơn vị là 1 và phần tử zero là 0 và $\varphi : A^*_\diamond \to M$ là một ánh xạ. Khi đó, φ được gọi là $\diamond$-đồng cấu vị nhóm nếu thỏa mãn các điều kiện sau:

(1) $\forall x, y \in A^*_\diamond$, $x.y \neq \theta$ thì $\varphi(x.y) = \varphi(x).\varphi(y)$

(2) $\varphi(e) = 1$

(3) $\varphi(\theta) = 0$

Định nghĩa 3.2. Cho $L \subseteq A^*_\diamond$ và vị nhóm M. Ta nói vị nhóm M thỏa L nếu tồn tại một $\diamond$-đồng cấu vị nhóm $\varphi : A^*_\diamond \to M$ sao cho $L = \varphi^{-1}(N)$, với $N \subseteq M$. Nếu $L = \varphi^{-1}(N)$ thì ta nói rằng L thỏa bởi φ.

Các kết quả sau là hiển nhiên, có thể kiểm tra trực tiếp từ định nghĩa.

Tính chất 3.1. Cho $N_1, N_2 \subseteq M$, khi đó

$$\varphi^{-1}(N_1 \cap N_2) = \varphi^{-1}(N_1) \cap \varphi^{-1}(N_2)$$
$$\varphi^{-1}(N_1 \cup N_2) = \varphi^{-1}(N_1) \cup \varphi^{-1}(N_2)$$
$$\varphi^{-1}(N_1 \backslash N_2) = \varphi^{-1}(N_1) \backslash \varphi^{-1}(N_2)$$

Hơn nữa, nếu φ là toàn ánh thì ta có

$$\varphi^{-1}(N_1^{-1} N_2) = \varphi^{-1}(N_1)^{-1} \varphi^{-1}(N_2)$$
$$\varphi^{-1}(N_1 N_2^{-1}) = \varphi^{-1}(N_1) \varphi^{-1}(N_2)^{-1}$$

Với mỗi $\diamond$-ngôn ngữ $L \subseteq A^*_\diamond$, áp dụng cách thức tương tự của S. Eilenberg [11] ta có thể xây dựng vị nhóm M thỏa L trên $A^*_\diamond$. Ký hiệu $\mathcal{R}(A, M)$ là tập tất cả các $\diamond$-ngôn ngữ trên $A^*_\diamond$ thỏa bởi M. Theo [9], $\mathcal{R}(A, M)$ đóng với các phép toán Boole. Hơn nữa, nếu φ là toàn cấu thì $\mathcal{R}(A, M)$ đóng với phép lấy thương trái và thương phải.

Kết hợp với một otomat đa định hữu hạn $\mathcal{A} = (A, Q, \delta, q_0, T)$, trong [8] đã định nghĩa một hình thức đặc biệt của otomat (gọi là $\diamond$-*otomat*) $\mathcal{A}_\diamond = (A_\diamond, Q_\diamond, \delta_\diamond, I_\diamond, T_\diamond)$ đoán nhận tập các $\diamond$-từ trên $A^*_\diamond$. Từ đó cho phép định nghĩa $\diamond$-đoán nhận được và thiết lập các kết quả cơ sở dựa trên $\diamond$-otomat, $\diamond$-đoán nhận được trên tập các từ định biên như sau:

Một dãy $x_1, ..., x_n, x_i \in A_\diamond$ được *đoán nhận* bởi $\diamond$-otomat $\mathcal{A}_\diamond$ khi và chỉ khi tồn tại một trạng thái $q_\diamond \in T_\diamond$, sao cho $q_\diamond \in \delta_\diamond(\delta_\diamond(\delta_\diamond(q_0, x_1), \ldots), x_n)$ và trong trường hợp này, $\diamond$-từ $x = x_1 \ldots x_n$ cũng được đoán nhận bởi $\diamond$-otomat $\mathcal{A}_\diamond$. Ký hiệu $\mathcal{L}(\mathcal{A}_\diamond)$ là tập tất cả các $\diamond$-từ được đoán nhận bởi $\diamond$-otomat $\mathcal{A}_\diamond$, ta có

$$\mathcal{L}(\mathcal{A}_\diamond) = \{\, x \in A^*_\diamond \mid \exists q_0 \in I_\diamond \textit{ sao cho } \delta_\diamond(q_0, x) \cap T_\diamond \neq \emptyset\}$$

Định nghĩa 3.3. Một $\diamond$-ngôn ngữ $L \subseteq A^*_\diamond$ được gọi là $\diamond$-đoán nhận được ($\diamond$-ngôn ngữ đoán nhận được) nếu $L = \mathcal{L}(\mathcal{A}_\diamond)$, với $\mathcal{A}_\diamond$ là $\diamond$-otomat nào đó.

Mối quan hệ không tầm thường giữa tính đoán nhận được trên $A^*_\diamond$ và trên A^* cho bởi mệnh đề sau:

Mệnh đề 3.1. *Cho $L \subseteq A^*$. Nếu L là đoán nhận được thì $Proj^{-1}(L)$ là $\diamond$-đoán nhận được. Tồn tại $L \subseteq A^*_\diamond$ không là $\diamond$-đoán nhận được nhưng $Proj(L)$ là đoán nhận được.*

Từ các định nghĩa và mệnh đề trên, ta có các kết quả sau:

Định lý 3.1. *Cho $L \subseteq A^*_\diamond$. Khi đó, các điều kiện sau đây là tương đương:*

- *(i) L được đoán nhận bởi một $\diamond$-otomat đa định hữu hạn.*
- *(ii) L được đoán nhận bởi một $\diamond$-otomat đơn định hữu hạn.*
- *(iii) L thỏa bởi một vị nhóm hữu hạn.*

4. MÃ VỚI TỪ ĐỊNH BIÊN VÀ THUẬT TOÁN

Trong phần này, ta nhắc lại khái niệm mã với các từ định biên (hay gọi là $\diamond$-*mã*) trên $A^*_\diamond$ và một số kết quả được trình bày trong [9]. Từ đó, thiết lập một số đặc trưng cho $\diamond$-mã, mã và mã luân phiên (Định lý 4.1) và đề xuất thuật toán kiểm định mã luận phiên mới theo tiếp cận thông qua thuật toán kiểm định $\diamond$-mã cùng với việc phân tích đánh giá độ phức tạp của thuật toán.

Định nghĩa 4.1. Cho $X \subseteq A^+_\diamond$. Khi đó, Xđược gọi là $\diamond$-mã nếu với mọi $x \neq \theta$ thuộc $A^+_\diamond$, x thừa nhận nhiều nhất một phân tích trong X.

Ví dụ 4.1. Cho $X = \{(0,a,1),(0,b,1),(1,c,0),(1,d,0)\}$. Khi đó, X là $\diamond$-mã.

Ví dụ 4.2. Cho $X = \{(0,a,1),(1,b,0),(0,ab,0),(0,ba,1)\}$, ta dễ dàng nhận thấy X không là $\diamond$-mã, vì $\diamond$-từ $x = (0,ababab,0)$ có hai phân tích trong X:

$$f_1 : \ (0,ab,0).(0,ab,0).(0,ab,0) \ ;$$
$$f_2 : \ (0,a,1).(1,b,0).(0,ab,0).(0,a,1).(1,b,0) \ .$$

Nhận xét 4.1. Các ví dụ sau cho ta thấy rằng, có X là $\diamond$-mã nhưng $Proj(X)$ không là mã và ngược lại, có X không là $\diamond$-mã trong khi $Proj(X)$ là mã.

Ví dụ 4.3. Cho $X = \{(0,a,1),(0,b,1),(1,aa,1)\}$ là $\diamond$-mã, nhưng $Proj(X) = \{a,b,aa\}$ không là mã.

Ví dụ 4.4. Cho bảng chữ $A = \{a,b\}$ và $X = \{(i,a,j) \mid a \in A, i,j \in B\}$ không là $\diamond$-mã, nhưng $Proj(X) = \{a,b\}$ là mã.

Định lý sau biểu diễn mối quan hệ phân lớp giữa mã, mã luân phiên với $\diamond$-mã, cho thấy rằng $\diamond$-mã, mã luân phiên được xem là hình thức mở rộng của mã.

Định lý 4.1. *Cho $C \subseteq A^+$. Khi đó, các điều kiện sau là tương đương:*

- *(i) C là mã.*
- *(ii) Cặp (C, C) là mã luân phiên.*
- *(iii) $(C_\lhd \cup C_\rhd)$ là $\diamond$-mã, với $C_\lhd = \{(0,w,1) \mid w \in C\}$, $C_\rhd = \{(1,w,0) \mid w \in C\}$.*

Chứng minh. $(i) \Rightarrow (ii)$. Ta chứng minh C là mã thì cặp (C, C) là mã luân phiên.

Từ giả thiết C là mã suy ra: Nếu từ $w \in A^+$ thừa nhận sự phân tích luân phiên $w_1 w_2 ... w_{2k}$ thì phân tích đó là duy nhất và đây cũng chính là một phân tích luân phiên theo (C, C). Vì vậy, cặp (C, C) là mã luân phiên.

$(ii) \Rightarrow (iii)$. Đặt $Z = (C_\lhd \cup C_\rhd)$, ta chứng minh cặp (C, C) là mã luân phiên thì Z là $\diamond$-mã.

Phản chứng, giả sử ngược lại Z không là $\diamond$-mã. Khi đó, tồn tại một $\diamond$-từ $x \neq \theta$, x thuộc $A^+_\diamond$ có hai phân tích khác nhau trong Z.

$$x = z_1.z_2 ... z_n = z'_1.z'_2 ... z'_m \text{ , với } z_1 \neq z'_1,\ m, n \geq 1,\ z_i, z'_j \in Z..$$

Trường hợp $z_1 \in C_\lhd$, thì $z_2 \in C_\rhd$, $z'_1 \in C_\lhd$, và nếu $z_{2k-1} \in C_\lhd$ thì $z_{2k} \in C_\rhd$, với $\forall k \geq 1$, và nếu $z'_{2l-1} \in C_\lhd$ thì $z'_{2l} \in C_\rhd$, với $l \geq 1$.

Do vậy, ta có $x_1.y_2 ... x_{n-1}.y_n = Proj(z_1.z_2 ... z_n) = Proj(z'_1.z'_2 ... z'_m) = x'_1.y'_2 ... x'_{m-1}.y'_m$. Dễ thấy rằng (C, C) không là mã luân phiên, mâu thuẫn. Chứng minh tương tự cho trường hợp $z_1 \in C_\rhd$. Do đó, $Z = (C_\lhd \cup C_\rhd)$ là $\diamond$-mã.

$(iii) \Rightarrow (i)$. Đặt $Z = (C_\lhd \cup C_\rhd)$, ta chứng minh rằng, nếu Z là $\diamond$-mã thì C là mã.

Phản chứng, giả sử ngược lại Z là $\diamond$-mã những C không là mã. Khi đó, tồn tại một từ $w \in A^+$ có hai phân tích khác nhau trong C :

$$w = w_1.w_2 ... w_n = w'_1.w'_2 ... w'_m \quad (w_1 \neq w'_1)$$

Từ biểu thức trên, suy ra:

$$w' = w_1.w_2 ... w_n.w_1.w_2 ...w_n = w'_1.w'_2 ... w'_m.w'_1.w'_2 ... w'_m$$

Bởi $2n$, $2m$ là số các chữ của hai vế là chẵn và w' có hai phân tích khác nhau trong C, hai phân tích này tương đương với hai phân tích luân phiên của một $\diamond$-từ trong Z :

$$(0, w_1, 1).(1, w_2, 0) \dots (1, w_n, 0) = (0, w'_1, 1).(1, w'_2, 0) \dots (1, w'_m, 0)$$

Nghĩa là, tồn tại một $\diamond$-từ có hai phân tích khác nhau trong Z, mâu thuẫn. Do đó, C là mã. ■

Trong trường hợp $Z = X \cup Y$ là $\diamond$-mã, với $X \subseteq A^+_{\lhd \setminus \varepsilon}$ và $Y \subseteq A^+_{\rhd \setminus \varepsilon}$ thì ta cũng gọi Z là *mã luân phiên trên* $A^*_\diamond$. Từ Định lý 4.1, ta nhân thấy, nếu Z là $\diamond$-mã thì $Proj(Z \cap A^*_\triangle)$, $Proj(Z \cap A^*_\triangledown)$ là mã và cặp $(Proj(Z \cap A^*_\lhd), Proj(Z \cap A^*_\rhd))$ là mã luân phiên.

Tương tự phương pháp của Sardinas-Patterson (xem [10]), ta thiết lập thủ tục sau để kiểm định một $\diamond$-ngôn ngữ cho trước có là $\diamond$-mã hay không.

Thủ tục kiểm định $\diamond$-mã

Input: *Cho* $X \subseteq A^+_\diamond$.

Output: Kết luận X *là* $\diamond$*-mã hoặc không.*

Bước_ 1. $U_1 = (X^{-1})X - e - \{(i, \varepsilon, i) \mid i \in B\}$, $n = 2$

Bước_ 2. (*Loop*)

$$U_n = (U_{n-1})^{-1}X \cup (X^{-1})U_{n-1}$$

Bước_ 3. *If* $e \in U_n$ *Then* *goto* *Bước_ 5*

If $U_n = \emptyset$ *or* $U_k = U_n \neq \emptyset$ (*for any* $k = 1, ..., n-1$) *Then goto Bước_4*
Else $n = n + 1$, *Loop*

Bước_4. Thông báo "X là $\diamond$*-mã" và Kết thúc.*

Bước_5. Thông báo "X không là $\diamond$*-mã" và Kết thúc.*

Tính đúng đắn của thủ tục này dựa trên Bổ đề và Định lý sau [9]:

Bổ đề 4.1. *Cho* $X \subseteq A_\diamond^+$ *và* U_n *được xác định như trên. Với* $\forall n \geq 1, k = 1, ..., n, e \in U_n$ *khi và chỉ khi tồn tại* $\diamond$*-từ* $z \in U_k$ *và* $d, l \geq 0$ *sao cho*

$$z.x_1.x_2 \dots x_d \;=\; x'_1.x'_2 \dots x'_l, \quad k = n - d - l,\ x_i, x'_j \in X$$

Định lý 4.2. *Cho* $X \subseteq A_\diamond^+$ *và* U_n *được xác định như trên. Khi đó,* X *là* $\diamond$*-mã khi và chỉ khi với mọi* $n \geq 1, e \notin U_n$.

Với X là $\diamond$-ngôn ngữ, trong trường hợp X là $\diamond$-đoán nhận được thì ta có kết quả sau:

Hệ quả 4.1. *Cho* $X \subseteq A_\diamond^+$ *là* $\diamond$*-đoán nhận được. Các bước kiểm tra X có là* $\diamond$*-mã hay không bởi thủ tục ở trên luôn luôn dừng sau một số hữu hạn bước lặp. Tồn tại thuật toán quyết định X có là* $\diamond$*-mã hay không.*

Từ Mệnh đề 3.1, Định lý 4.1 và Hệ quả 4.1, ta nhận được kết quả sau:

Hệ quả 4.2. *Cho* $X, Y \subseteq A^+$ *là đoán nhận được. Tồn tại thuật toán để quyết định cặp* (X, Y) *là mã luân phiên hay không có độ phức tạp thời gian cỡ* $O(2^{|M|})$*, với M là một vị nhóm hữu hạn thỏa đồng thời cả* $X_\lhd$ *và* $Y_\rhd$.

Chứng minh. Giả sử $Z = (X_\lhd \cup Y_\rhd)$ là $\diamond$-đoán nhận được, ta có thể xây dựng một $\diamond$-otomat đoán nhận tất cả các tập Z, $\{e\}$ và $\{(i, \varepsilon, i) \mid i \in B\}$. Dựa vào Định lý 3.1, ta có thể xây dựng một $\diamond$-toàn cấu $\varphi : A_\diamond^* \to M$, M hữu hạn, sao cho φ thỏa Z, $\{e\}$, $\{(i, \varepsilon, i) \mid i \in B\}$.

Sử dụng Tính chất 3.1: φ^{-1} bảo toàn với các phép toán Boole, phép lấy thương trái, thương phải trên các tập con của M. Ta có thể kết luận tất cả các tập $U_{n\geq 1}$ được định nghĩa ở trên thỏa bởi φ. Vì M là hữu hạn, nên thủ tục ở trên sẽ dừng sau không quá $2^{|M|}$ bước lặp. Vì vậy, độ phức tạp thời gian của thuật toán kiểm định mã luân phiên mới trong trường hợp xấu nhất là $O(2^{|M|})$. ■

Trước khi đánh giá hai thuật toán kiểm định mã luân phiên, ta thiết lập định lý sau: Cho M là một vị nhóm với đơn vị là 1_M. Đặt

$$M_\diamond = \{(i, m, j) \mid\ m \in M, i, j \in B\} \cup \{\mathbf{0}, \mathbf{1}\},$$
với $\mathbf{0}$, $\mathbf{1} \notin M$ và $\mathbf{0}$ đóng vai trò zero của $M_\diamond$, $\mathbf{1}$ đóng vai trò đơn vị của $M_\diamond$.

$$M_\lhd = \{(0, m, 1) \mid\ m \in M\}\ ; \qquad M_\rhd = \{(1, m, 0) \mid\ m \in M\}\ ;$$
$$M_\triangle = \{(0, m, 0) \mid\ m \in M\}\ ; \qquad M_\triangledown = \{(1, m, 1) \mid\ m \in M\}\ .$$

Định lý 4.3. *Cho* $X \subseteq A^*$ *và đồng cấu vị nhóm* $\alpha : A^* \to M$ *thỏa X. Xét* $\diamond$*-đồng cấu vị nhóm* $\alpha_\diamond : A_\diamond^* \to M_\diamond$ *cho bởi*

(1) $\forall x, y \in A_\diamond^*,\ x.y \neq \theta$ *thì* $\alpha_\diamond(x.y) = \alpha_\diamond(x).\alpha_\diamond(y)$

(2) $\alpha_\diamond(\theta) = \mathbf{0}$.

(3) $\alpha_\diamond(e) = \mathbf{1}$.

(4) $\alpha_\diamond(i, \varepsilon, j) = (i, 1_M, j)$.

Khi đó, $\alpha_\diamond$ thỏa đồng thời $A^*_\diamond$, $A^*_\lhd$, $A^*_\rhd$, $A^*_\triangle$, $A^*_\triangledown$, $X_\diamond$, $X_\lhd$, $X_\rhd$, $X_\triangle$, $X_\triangledown$.

Chứng minh. Theo giả thiết $\alpha : A^* \to M$ thỏa X, nghĩa là $X = \alpha^{-1}(N)$, với $N \subseteq M$. Khi đó, ta có

$$\alpha_\diamond^{-1}(M_\diamond) = A^*_\diamond;\ \ \alpha_\diamond^{-1}(M_\lhd) = A^*_\lhd;\ \ \alpha_\diamond^{-1}(M_\rhd) = A^*_\rhd;\ \ \alpha_\diamond^{-1}(M_\triangle) = A^*_\triangle;\ \ \alpha_\diamond^{-1}(M_\triangledown) = A^*_\triangledown.$$

Đặt $X' = \{(i, w, j) \in A^*_\diamond \mid w \in X, i, j \in B\}$ và $X_\diamond = X'$ nếu $\varepsilon \notin X$; $X_\diamond = X' \cup \{e\}$ nếu $\varepsilon \in X$. Khi đó sẽ tồn tại $N_\diamond \subseteq M_\diamond$ sao cho $X_\diamond = \alpha_\diamond^{-1}(N_\diamond)$. Do đó

$$\begin{aligned}
X_\lhd &= X_\diamond \cap A^*_\lhd = \alpha_\diamond^{-1}(N_\diamond) \cap \alpha_\diamond^{-1}(M_\lhd);\\
X_\rhd &= X_\diamond \cap A^*_\rhd = \alpha_\diamond^{-1}(N_\diamond) \cap \alpha_\diamond^{-1}(M_\rhd);\\
X_\triangle &= X_\diamond \cap A^*_\triangle = \alpha_\diamond^{-1}(N_\diamond) \cap \alpha_\diamond^{-1}(M_\triangle);\\
X_\triangledown &= X_\diamond \cap A^*_\triangledown = \alpha_\diamond^{-1}(N_\diamond) \cap \alpha_\diamond^{-1}(M_\triangledown).
\end{aligned}$$

Từ đó, suy ra $A^*_\diamond$, $A^*_\lhd$, $A^*_\rhd$, $A^*_\triangle$, $A^*_\triangledown$, $X_\diamond$, $X_\lhd$, $X_\rhd$, $X_\triangle$, $X_\triangledown$ thỏa bởi $\alpha_\diamond$. ■

*** Độ phức tạp của thuật toán kiểm định mã luân phiên qua các tiếp cận**

Cho X, Y là các ngôn ngữ chính quy được thỏa bởi các đồng cấu vị nhóm $\alpha : A^* \to M$ và $\beta : A^* \to N$ tương ứng. Khi đó, bước (B_1 - Kiểm tra $Z = XY$ là mã) của thuật toán kiểm định mã luân phiên trong [7] có độ phức tạp cỡ $O(2^{2^{(|M|+|N|)}})$.

Mặt khác, xuất phát từ tiếp cận thông qua thuật toán kiểm định $\diamond$-mã và Định lý 4.3, suy ra $\alpha_\diamond : A^*_\diamond \to M_\diamond$ thỏa $X_\lhd$ và $\beta_\diamond : A^*_\diamond \to N_\diamond$ thỏa $Y_\rhd$, với $|M_\diamond| = 4|M| + 2$, $|N_\diamond| = 4|N| + 2$. Do đó, ta có thể xây dựng $\diamond$-toàn cấu $\varphi_\diamond = \alpha_\diamond \circ \beta_\diamond$ thỏa $(X_\lhd \cup Y_\rhd)$ là $\varphi_\diamond : A^*_\diamond \to M_\diamond \times N_\diamond$. Khi đó, thuật toán kiểm định (X, Y) là mã luân phiên (tương ứng với thuật toán kiểm định $Z = X_\lhd \cup Y_\rhd$ là $\diamond$-mã) có độ phức tạp trong trường hợp tồi nhất (theo Hệ quả 4.2) cỡ $O(2^{|M_\diamond \times N_\diamond|}) = O(2^{(4|M|+2).(4|N|+2)})$.

Với $|M|$, $|N|$ đủ lớn, dễ thấy $2^{(4|M|+2).(4|N|+2)} < 2^{2^{(|M|+|N|)}}$. Vì vậy, nếu chỉ thực hiện bước (B_1) của thuật toán kiểm định mã luân phiên trong [7], thì đã có độ phức tạp cỡ $O(2^{2^{(|M|+|N|)}})$. Trong khi đó, nếu tiếp cận thông qua thuật toán kiểm định $\diamond$-mã, thì thuật toán kiểm định mã luân phiên mới có độ phức tạp giảm xuống chỉ còn cỡ $O(2^{(4|M|+2).(4|N|+2)})$, với M, N là các vị nhóm hữu hạn thỏa X, Y tương ứng. Điều này cho thấy một phần khả năng nghiên cứu về $\diamond$-ngôn ngữ và $\diamond$-mã.

5. KẾT LUẬN

Bài báo đưa ra một hệ mã mới - mã của các từ định biên và thiết lập thuật toán kiểm định mã của các từ định biên trên các $\diamond$-ngôn ngữ đoán nhận được. Thông qua thuật toán này, cho phép xây dựng một thuật toán mới để kiểm định mã luân phiên [7], làm giảm rõ rệt về độ phức tạp. Một sơ đồ quan hệ phân lớp giữa ba loại mã - mã, mã luân phiên và $\diamond$-mã được đề xuất, cho thấy $\diamond$-mã như là một sự mở rộng của mã và mã luân phiên. Nghiên cứu về tính chính quy, độ trễ giải mã, làm đầy mã ... của $\diamond$-ngôn ngữ là những hướng lý thú, sẽ được nghiên cứu trong các công trình tiếp theo.

TÀI LIỆU THAM KHẢO

[1] M. P. Schützenberger, On a question concerning certain free submonoids, *Journal of Combinatorial Theory* **1** (4) (1966) 437–442.

[2] Jean-Eric Pin, "Variété des Languages Infinis et variete de semigroupes", These Docteur d'Etat (1982).

[3] Jean-Eric Pin, Pascal Weil, Polynomial closure and unambiguous products. *Theory of Computing Systems* **30** (1977) 383–422.

[4] Pascal Weil, Groups, Codes and unambiguous automata. *Theoretical aspects of computer science, 2nd ann. Symp.*, Saarbrcken/Ger. 1985, Lect. Notes Comput. Sci. **Vol. 182** (1985) 351-362.

[5] Phan Trung Huy, Do Long Van, On Non-Ambiguous Büchi V-automata. *Proceedings of the Third Asian Mathematical Conference*, Philippines, Oct., 23-27, 2000 (World Scientific 2002).

[6] Phan Trung Huy, Vũ Thành Nam, Mã luân phiên và mã tiền ngữ cảnh, *Kỷ yếu Hội thảo quốc gia lần thứ VII, Một số vấn đề chọn lọc của công nghệ thông tin*, Đà Nẵng, 8/2004 (188–197).

[7] Ho Ngoc Vinh, Vu Thanh Nam, Phan Trung Huy, Codes base on unambiguous products, *Proceedings of the 2nd International Conference on Computational Collective Intelligence - Technologies and Applications (ICCCI 2010)*, Lectures Notes on Artificial Intelligence **Vol. 6423** (2010), 252-262. Springer-Verlag Berlin Heidelberg 2010.

[8] Hồ Ngọc Vinh, Vũ Thành Nam, Phan Trung Huy, Mã với các hình thức tích mới, *Kỷ yếu Hội thảo quốc gia lần thứ XII, Một số vấn đề chọn lọc của công nghệ thông tin và truyền thông*, Biên Hòa, 5-6/8, 2009 (186–197).

[9] Ho Ngoc Vinh, Phan Trung Huy, Codes of bounded words, *Proceedings of the 3rd International Conference on Computer and Electrical Engineering (ICCEE 2010)* **Vol. 2** (2010) 89–95. IEEE 2010.

[10] Jean Berstel, Dominique Perrin, *Theory of Codes*, Academic Press Inc., NewYork, 1985.

[11] Samuel Eilenberg, *Automata, languages and machines*, Vol. B, Academic Press, New York , 1976.

Nhận bài ngày 16 - 9 - 2010
Nhận lại sau sửa ngày 15 - 11 - 2010