Tạp chí Tin học và Điều khiển học, T. 18, S. 1 (2002), 80-86

CÂY SINH SỐ
ĐOÁN NHẬN TÍNH ĐỒNG DƯ, SỐ NGUYÊN TỐ
VÀ XÁC ĐỊNH ĐỘ PHỨC TẠP CỦA NÓ

VŨ TRỌNG QUẾ

Abstract. The article is on the graph tree labelled with natural digits deriving natural numbers no more than 10^m, in which m is a natural number. The graph tree shows the divisibility, prime number and its complexity.

Tóm tắt. Bài này trình bày thuật toán xây dựng cây sinh số đoán nhận tính đồng dư của các số tự nhiên nhỏ hơn 10^m với m là số tự nhiên dương hữu hạn tùy ý.

1. MỞ ĐẦU

Đồ thị là ngành khoa học được phát triển từ lâu và có nhiều ứng dụng hiện đại. Nhiều ý tưởng cơ bản của nó được đưa ra từ thế kỷ 18 bởi nhà toán học Thụy sỹ Leonhard Euler. Ông đã dùng đồ thị để giải quyết bài toán cầu Konigberg nổi tiếng. Thực tế đồ thị được áp dụng để giải quyết nhiều loại bài toán trong các lĩnh vực khoa học khác nhau như Vật lý, Hóa học... và nhiều lĩnh vực trong đời sống xã hội như xây dựng, giao thông vận tải, truyền thông,... Đặc biệt từ khi Tin học ra đời, việc dùng đồ thị để giải quyết các bài toán trên máy tính được thuận lợi và nhanh gọn hơn. Bài báo này trình bày thuật toán xây dựng cây sinh số đoán nhận tính đồng dư và số nguyên tố. Đây là một minh họa về ứng dụng của lí thuyết đồ thị trong số học.

2. ĐỒ THỊ SINH SỐ

Đồ thị có hướng (có thể có khuyên) G tách ra một đỉnh được gọi là đỉnh vào, đỉnh xuất phát hay đỉnh gốc (và đặt trong ô tròn có mũi tên), một tập con các đỉnh được gọi là các đỉnh ra hay đỉnh kết (mỗi đỉnh kết được đặt trong một ô chữ nhật), các đỉnh còn lại được gọi là đỉnh không kết (mỗi đỉnh được đặt trong một ô tròn) đồng thời mỗi cung t được ghi một chữ số thập phân a, ($a \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$) được gọi là đồ thị sinh số. Kí hiệu a được gọi là nhãn của cung t.

Để thu gọn cách biểu diễn người ta qui ước như sau: nếu từ đỉnh x sang đỉnh y có nhiều cung thì từ x sang y chỉ vẽ một cung và trên đó ghi đầy đủ các nhãn thuộc các cung đi từ x sang y.

Giả sử $D = t_1 t_2 \ldots t_m$ là một đường đi trong đồ thị sinh số G và a_i là nhãn của cung t_i $(1 \leq i \leq m)$. Dãy $\overline{D} = a_1 a_2 \ldots a_m$ được gọi là số sinh bởi đường D.

Tập gồm tất cả các số, mà mỗi số này được sinh bởi một đường trong G xuất phát từ đỉnh vào và đi tới một trong những đỉnh kết được gọi là tập số sinh bởi đồ thị G.

Thuật toán xây dựng đồ thị sinh số đoán nhận tính đồng dư

Với mỗi số nguyên dương m $(m \geq 2)$ ta có thể xây dựng đồ thị G sinh tất cả các số tự nhiên chia hết cho m. Thuật toán được xây dựng như sau:

Bước 1: xây dựng đỉnh vào, đỉnh kết ghi số 0 và $m-1$ đỉnh không kết ghi các số từ 1 đến $m-1$;

Bước 2: gán giá trị $x = 1$;

Bước 3: nếu $x \leq 9$ thực hiện bước 4,
nếu không thực hiện bước 9;

Bước 4: gán r = số dư của phép chia x cho m;

Bước 5: nếu chưa có cung đi từ đỉnh gốc đến đỉnh ghi số r thực hiện bước 6, nếu không thực hiện bước 7;

Bước 6: xây dựng cung đi từ đỉnh gốc đến đỉnh r;

Bước 7: gắn nhãn hoặc thêm nhãn x cho cung đi từ đỉnh gốc đến đỉnh r;

Bước 8: gán $x = x + 1$, quay lại thực hiện bước 3;

Bước 9: gán $i = 0$;

Bước 10: nếu $i \leq m-1$ thực hiện bước 11, nếu không thì thực hiện bước 17;

Bước 11: gán $j = 0$;

Bước 12: nếu $j \leq 9$ thực hiện bước 13, nếu không thì thực hiện bước 16;

Bước 13: gán r bằng số dư của phép chia $(10 * i + j)$ cho m;

Bước 14: xây dựng cung từ đỉnh ghi số i đến đỉnh ghi số r và số gắn cho cung này nhãn j;

Bước 15: gán $j = j + 1$; quay lại thực hiện bước 12;

Bước 16: gán $i = i + 1$; quay lại thực hiện bước 10;

Bước 17: thu gọn cung;

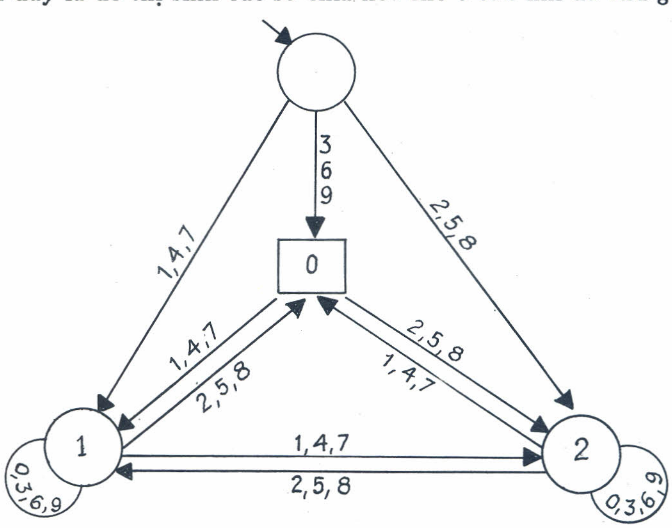
Bước 18: in kết quả.

Dễ dàng nhận thấy rằng, $\forall r$ $(0 \leq r \leq m-1)$ đỉnh r sinh tất cả các số nguyên dương đồng dư với r theo mod m.

Nếu trong đồ thị sinh số, ta không quan tâm đến số được sinh chia cho m dư bao nhiêu, chỉ quan tâm nó chia hết cho m hay không, ta bỏ các số cụ thể ghi trên đỉnh của đồ thị và thay thế đỉnh cũ bằng đỉnh mới là đỉnh kết hay không kết. (Nếu ta xây dựng đồ thị bằng hình vẽ thì đỉnh kết là hình vuông, đỉnh không kết là hình tròn).

Ví dụ: Nếu mục đích của bài toán vẽ đồ thị sinh các số chia hết cho m, ta sẽ vẽ đỉnh ghi số 0 là đỉnh hình vuông, tất cả các đỉnh còn lại (trừ đỉnh gốc) là đỉnh hình tròn. Ngược lại, nếu mục đích của bài toán là xây dựng đồ thị sinh các số không chia hết cho m ta sẽ vẽ đỉnh ghi số 0 là đỉnh hình tròn, tất cả các đỉnh còn lại (trừ đỉnh gốc) là đỉnh hình vuông.

Ví dụ đồ thị sau đây là đồ thị sinh các số chia hết cho 3 sau khi đã thu gọn cung.



Việc xây dựng đồ thị sinh số, ở đây không đánh giá độ phức tạp của thuật toán trên theo góc độ thời gian và không gian, chỉ xem số đỉnh cần tối thiểu để đồ thị sinh các số theo yêu cầu của bài toán là bao nhiêu. Số đỉnh của đồ thị được gọi là độ phức tạp của đồ thị. (Đồ thị sinh các số chia hết cho 3 có độ phức tạp là 4).

Cây sinh số đoán nhận tính đồng dư

Đồ thị sinh số nguyên dương hữu hạn G được gọi là cây sinh số, nếu nó có ít nhất hai đỉnh và thỏa mãn đồng thời 3 điều kiện sau:

1) Mỗi đỉnh khác đỉnh vào là đỉnh cuối của một cung duy nhất.

2) Đỉnh vào $\delta(G)$ không là đỉnh cuối của bất kỳ cung nào.

3) Đồ thị G không có vòng.

Đỉnh vào của cây sinh số được gọi là gốc của cây.

Đối với một cây sinh số ta còn phân tầng theo nguyên tắc sau: Đỉnh gốc được thừa nhận là đỉnh tầng 0. Đỉnh có cung đi tới từ đỉnh gốc là đỉnh tầng 1. Đỉnh có cung đi từ đỉnh tầng i là đỉnh tầng $i+1$. Số đỉnh của cây gọi là độ phức tạp của cây.

Từ đồ thị sinh số đoán nhận tính đồng dư ta có thể tạo ra cây sinh số nguyên dương hữu hạn nhỏ hơn 10^m đoán nhận tính đồng dư bằng thuật toán sau:

Bước 1: xây dựng đỉnh gốc;
Bước 2: gán giá trị $x = 1$;
Bước 3: đánh dấu đã thăm đỉnh gốc;
Bước 4: nếu $x \le n$ thực hiện bước 5,
nếu không thực hiện bước 12;
Bước 5: gán giá trị q = bằng độ dài của x; (ví dụ nếu $x = a_1a_2\dots a_m$, $q = m$)
Bước 6: gán giá trị $i = 1$;
Bước 7: nếu $i \le q$ thực hiện bước 8,
nếu không thực hiện bước 11;
Bước 8: gán giá trị $t = a_i$;
Bước 9: nếu cung từ đỉnh đã đánh dấu ở tầng $i-1$ đến đỉnh ở tầng i gắn nhãn t chưa được xây dựng, thực hiện bước 10,
nếu không ta thực hiện đánh dấu đỉnh ở tầng i, đỉnh đánh dấu mới này là đỉnh cuối của cung đi từ đỉnh đánh dấu cũ gắn nhãn t; gán $i = i+1$; quay lại thực hiện bước 7;
Bước 10: thêm đỉnh mới ở tầng i, đỉnh mới này cùng loại (kết hoặc không kết) với đỉnh cuối cùng của đường đi trên đồ thị sinh số, sinh ra số x đến vị trí i; xây dựng cung đi từ đỉnh đã đánh dấu ở tầng $i-1$ đến đỉnh mới ở tầng i; gắn nhãn t cho cung mới xây dựng; đánh dấu đỉnh vừa mới xây dựng; gán giá trị $i = i+1$, quay lại bước 7;
Bước 11: gán giá trị $x = x+1$, bỏ đánh dấu ở đỉnh cũ, quay lại bước 3;
Bước 12: cho ra kết quả.

Tập số nguyên dương $\le 10^m$ là số hữu hạn. Cách xây dựng cây sinh số nguyên dương hữu hạn trên chứa đựng tính đầy đủ, đúng đắn và tính kết thúc của thuật toán. Do vậy ta khẳng định: Từ đồ thị sinh số đoán nhận tính đồng dư ta sẽ xây dựng được cây sinh số nguyên dương hữu hạn đoán nhận tính đồng dư.

Thuật toán xây dựng cây sinh số giao, đoán nhận tính đồng dư

Ta xây dựng G là cây sinh số giao của hai cây sinh số G_1 và G_2 như sau:

Bước 1: xây dựng đỉnh gốc;
Bước 2: gán giá trị $x = 1$;
Bước 3: đánh dấu đã thăm đỉnh gốc;
Bước 4: nếu $x \le n$ thực hiện bước 5,
nếu không thực hiện bước 12;

Bước 5: gán giá trị q bằng độ dài của x; (ví dụ nếu $x = a_1a_2 \ldots a_m$, $q = m$)

Bước 6: gán giá trị $i = 1$;

Bước 7: nếu $i \leq q$ thực hiện bước 8,
nếu không thực hiện bước 11;

Bước 8: gán giá trị $t = a_i$;

Bước 9: nếu cung từ đỉnh đã đánh dấu ở tầng $i-1$ đến đỉnh ở tầng i gắn nhãn t chưa được xây dựng thì thực hiện bước 10,
nếu không ta thực hiện đánh dấu đỉnh ở tầng i, đỉnh đánh dấu mới này là đỉnh cuối của cung đi từ đỉnh đánh dấu cũ gắn nhãn t; gán $i = i + 1$ quay lại thực hiện bước 7;

Bước 10: nếu $\exists\, a_1a_2 \ldots a_i \in G_1$ và $a_1a_2 \ldots a_i \in G_2$, thêm đỉnh mới ở tầng i, đỉnh mới này là đỉnh kết khi và chỉ khi đỉnh cuối cùng của đường đi trên đồ thị sinh số G_1 và G_2 sinh ra số x đến vị trí i đều là đỉnh kết; xây dựng cung đi từ đỉnh đã đánh dấu ở tầng $i-1$ đến đỉnh mới ở tầng i, gắn nhãn t cho cung mới xây dựng; đánh dấu đỉnh vừa mới xây dựng; gán giá trị $i = i + 1$, quay lại thực hiện bước 7,
nếu không thì thực hiện bước 11;

Bước 11: gán giá trị $x = x + 1$, bỏ đánh dấu ở đỉnh cũ, quay lại bước 3;

Bước 12: cho ra kết quả.

Ý nghĩa: Đồ thị G được xây dựng như trên là đồ thị sinh các số có cả hai tính chất mà đồ thị G_1 và G_2 có.

Ví dụ. Đồ thị G_1 sinh các số chia hết cho 2, đồ thị G_2 sinh các số chia hết cho 3 thì đồ thị G sinh các số đồng thời chia hết cho 2 và 3, hay đồ thị G sinh các số chia hết cho 6.

3. CÂY SINH SỐ ĐOÁN NHẬN TÍNH ĐỒNG DƯ VỚI MỘT TẬP HỢP SỐ BẤT KỲ

Với số tự nhiên m ($m \geq 1$) hữu hạn tùy ý, bằng các định nghĩa và thuật toán đã trình bày ở Phần 2, ta hoàn toàn có thể xây dựng được các cây sinh số đoán nhận tính đồng dư với bất kỳ số tự nhiên nào.

Để có cây sinh các số chia hết đồng thời cho 3 số m, n, p trước hết ta xây dựng cây G_m sinh các số chia hết cho m, cây G_n sinh các số chia hết cho n, cây G_p sinh các số chia hết cho p. Giao của ba cây trên là cây sinh các số chia hết đồng thời cho m, n, p.

Muốn có cây sinh các số không chia hết đồng thời cho 3 số m, n, p ta xây dựng cây H_m sinh các số không chia hết cho m, cây H_n sinh các số không chia hết cho n, cây H_p sinh các số không chia hết cho p. Giao của 3 cây trên là cây sinh các số không chia hết đồng thời cho m, n, p.

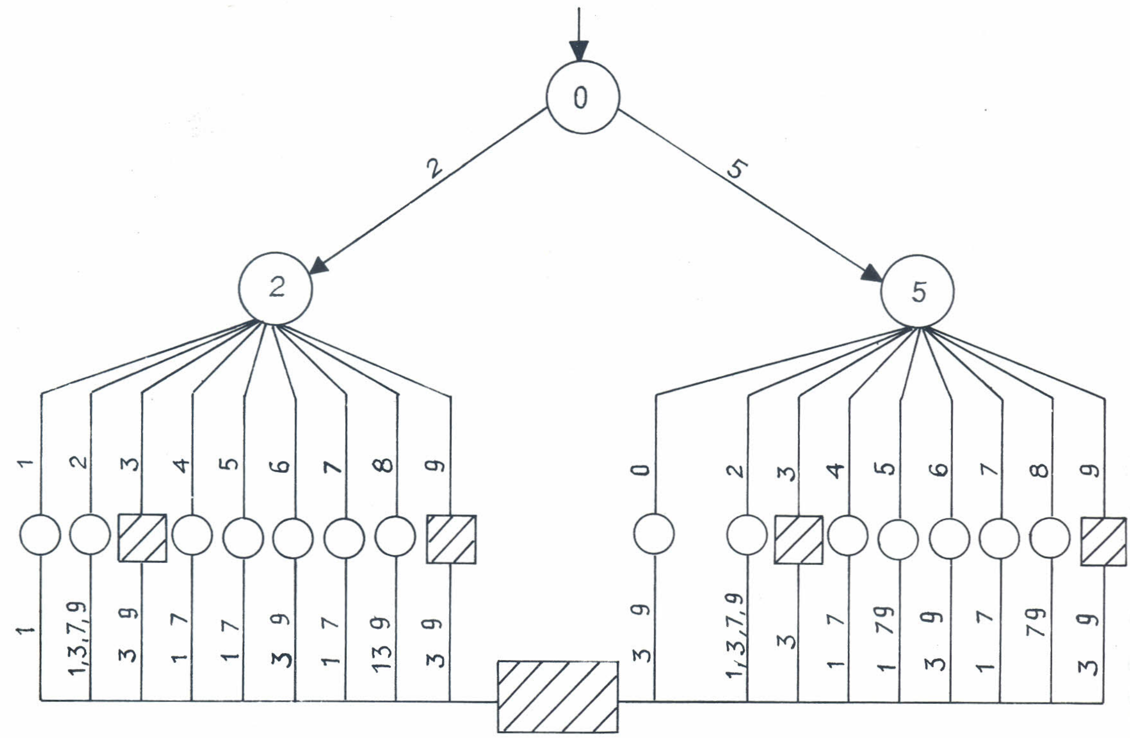
Ví dụ: Cây sinh số G_2 sinh các số không chia hết cho 2, cây sinh số G_3 sinh các số không chia hết cho 3. Giao của G_2, G_3 là cây sinh các số không chia hết cho 2 và cho 3.

Ví dụ ở hình vẽ 1 là cây sinh số thu gọn, sinh các số không chia hết đồng thời cho 2, 3, 5, 7, 11. (Do khổ giấy có hạn hình vẽ chỉ minh họa 2 nhánh tiêu biểu của cây mà có cung xuất phát từ đỉnh gốc gắn nhãn là 2 và 5).

4. THUẬT TOÁN XÂY DỰNG CÂY SINH CÁC SỐ NGUYÊN TỐ NHỎ HƠN $n = 10^m$

Ta biết rằng số tự nhiên $p > 1$ mà không có ước nào $\leq$ căn của p thì nó là số nguyên tố. Dựa trên nhận xét này, đối với mọi số tự nhiên $n = 10^m$ (m hữu hạn ≥ 1) ta có thể xây dựng cây đoán nhận các số nguyên tố nhỏ hơn n bằng thuật toán sau.

Khi xây dựng cây G_p (p nguyên tố tùy ý) đoán nhận các số không chia hết cho p, ta thừa nhận đỉnh sinh số p cũng là đỉnh kết của G_p (đỉnh kết sinh số p được gọi là đỉnh kết thừa nhận, các đỉnh kết còn lại được gọi là đỉnh kết thực sự).



Hình 1. Cây sinh số $< 10^3$ không chia hết cho 2, 3, 5, 7, 11.

Bước 1: gán giá trị k = phần nguyên của căn n+1;

Bước 2: gán **ketqua** = cây G_2;

Bước 3: $i = 3$;

Bước 4: nếu $i \leq k$, thực hiện bước 5,
nêu không ta thực hiện bước 8;

Bước 5: xây dựng cây G_i;

Bước 6: gán ketqua = ketqua $\cap$ G_i;

Bước 7: gán i = giá trị đỉnh kết thực sự sinh số nhỏ nhất của cây **ketqua**, quay lại thực hiện bước 4;

Bước 8: in kết quả.

Giả sử các thủ tục và hàm sau đây đã được xây dựng:

1/ Thủ tục xaydung(i) cho kết quả cây sinh số G_i sinh các số nhỏ hơn n và không chia hế cho i.

2/ Thủ tục giao(x, y) cho kết quả cây sinh số *ketqua* là giao của 2 cây x và y.

3/ Thủ tục timsotiep(x), cho kết quả số i có giá trị nhỏ nhất do đỉnh kết thực sự trong cây sinh ra.

4/ Thủ tục thugon(x), thủ tục thu gọn cung và đỉnh của cây sinh số x.

Ta có thể phác họa thuật toán trên theo ngôn ngữ lập trình Turbo Pascal như sau:

```
procedure songuyento(n):
      { khai báo các biến }
      { khai báo units }
begin
      k := int(sqrt(n)) + 1;
```

```
    ketqua := xaydung(2);
    i := 3;
    while i ≤ k do
        begin
            Xaydung(i);
            giao(ketqua,G_i);
            timsotiep(ketqua);
        end;
    thugon(ketqua);
    in ketqua;
end.
```

5. ĐỘ PHỨC TẠP CỦA CÂY SINH SỐ

Giả sử ta đã xây dựng được cây C thu gọn sinh các số nguyên tố $< 10^m$. Khi đó số tầng của C sẽ $\leq m+1$, được đánh số từ 0 đến m, số đỉnh (độ phức tạp) của cây sẽ là:

Số đỉnh ở tầng 0 là 1.

Số đỉnh ở tầng 1 không vượt quá 9.

Số đỉnh ở tầng $m-2$ sẽ không vượt quá $9 \times 10^{m-2}$.

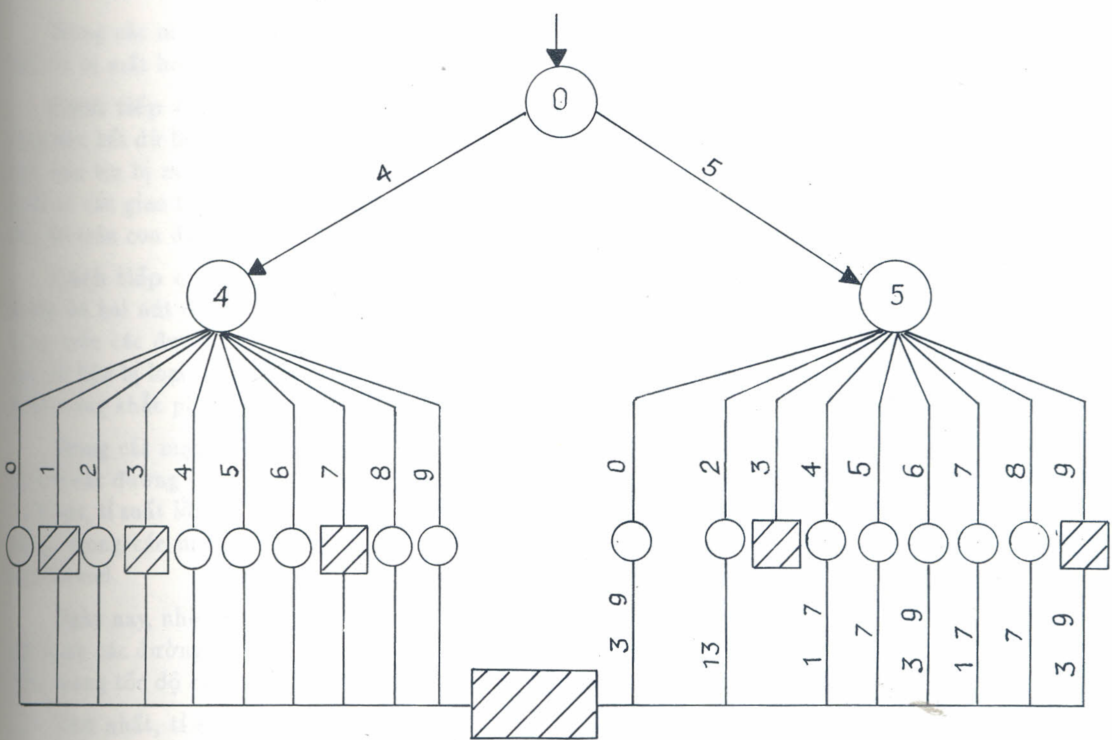
Số đỉnh ở tầng $m-1$ sẽ không vượt quá 30 đỉnh vì:

Số tận cùng của các nguyên tố lớn hơn 5 (ta coi trường hợp xây dựng cây sinh các số nguyên tố nhỏ hơn 5 là tầm thường) chỉ có thể là $\{1,3,7,9\}$ gồm 4 phần tử nên số các cung đi đến đỉnh ở tầng cuối sau khi dùng cách thu gọn cung có tối đa là $C_4^1 + C_4^2 + C_4^3 + C_4^4 = 15$ đỉnh kết và $C_4^1 + C_4^2 + C_4^3 + C_4^4 = 15$ đỉnh không kết.

Số đỉnh ở tầng cuối là 1.

Vậy ta có số đỉnh của cây đoán nhận các số nguyên tố nhỏ hơn 10^m không vượt quá:

$1+9+\ldots+9*10^{m-2}+2\times(2^4-1)+1 = 9(1+\ldots+(.10^{m-2}))+32 = 10^{m-1}-1+32 = 10^{m-1}+31.$



Hình 2. Cây sinh số nguyên tố < 1000. Độ phức tạp: 37

Ví dụ: Để xây dựng cây C_3 sinh các số nguyên tố < 1000 ($m = 3$) bằng phương pháp trên ta có độ phức tạp của cây nhỏ hơn $10^{m-1} + 31 = 131$.

Tuy về nguyên tắc cây C_3 có ước lượng độ phức tạp tới 131, song qua thực tế thì độ phức tạp của nó chỉ là 37. (Do khổ giấy có hạn hình 2 chỉ minh họa 2 nhánh tiêu biểu có cung xuất phát từ đỉnh gốc gắn nhãn là 4 và 5).

Tác giả xin chân thành cảm ơn PGS TS Đặng Huy Ruận về sự quan tâm hướng dẫn của Thày trong công việc nghiên cứu.

TÀI LIỆU THAM KHẢO

[1] Đặng Quang Ngàn, Luận văn Thạc sỹ, 1994.
[2] Đặng Huy Ruận: Độ phức tạp ôtômat hữu hạn của các dãy biểu thức chính qui suy rộng, *Tạp chí Khoa học Đại học Quốc gia Hà Nội* (1995).
[3] Đặng Huy Ruận, Giáo trình "Lý thuyết ngôn ngữ hình thức và Ôtômat", Trường Đại học Khoa học tự nhiên - ĐHQG Hà Nội.
[4 I. M. Vinogradov, *Cơ sở lý thuyết số* (tiếng Nga), Moskva, 1981.
[5] Kenneth H. Rosen, *Toán học rời rạc ứng dụng trong tin học*, NXB Khoa học và Kỹ thuật, Hà Nội, 1998.
[6] Nguyễn Văn Ba, *Ngôn ngữ hình thức*, Trường Đại học Bách khoa Hà Nội, 1997.
[7] Saloma A., *Nhập môn Tin học, Lý thuyết tính toán và các Ôtômat*, NXB Khoa học Kỹ thuật, Hà Nội, 1992.

Nhận bài ngày 10 - 9 - 200
Nhận lại sau khi sửa ngày 12 - 12 - 200

Khoa Toán - Cơ - Tin học,
Trường Đại học Khoa học tự nhiên,
Đại học Quốc gia Hà Nội.