TÌM HIỂU SÂU THÊM TOÁN HỌC SƠ CẤP

CÁC BÀI TOÁN NỔI TIẾNG VỀ DÃY CATALAN

NGÔ THỊ NHÃ - NGUYỄN VĂN LỢI
(ĐHTH Budapest, Hungary)

Bài viết này trình bày những bài toán đặc trưng về dãy số Catalan, chỉ ra mối quan hệ mật thiết của các bài toán, đưa ra một cách nhìn mới thông qua Bài toán 3.6 có tên gọi *kiến hành quân*. Với cách nhìn mới này, việc phát biểu bài toán Catalan trở nên rõ ràng hơn và bài toán Catalan tổng quát (cho nhiều chiều) được trình bày một cách đơn giản. Chúng tôi hy vọng đóng góp một cách tiếp cận mới cho các nghiên cứu mở rộng đề tài này.

1. Dãy Catalan

Các số Catalan (hay còn gọi là dãy Catalan) lần đầu tiên được Leonard Euler (1707 – 1783) quan tâm đến khi ông nghiên cứu vấn đề: có bao nhiêu cách có thể chia một đa giác thành các tam giác. Nhưng tên của dãy này lại thuộc về Eugene Charles Catalan (1814 – 1894) – một nhà toán học người Bỉ khi ông giải quyết thành công bài toán: có bao nhiêu cách để đóng ngoặc và mở ngoặc một dãy số khi thực hiện phép tính. Năm 1838, Catalan đã phát hiện ra rằng các số này là lời giải chung của rất nhiều bài toán tưởng chừng xa lạ. Có những bài để ở dạng này thì vô cùng phức tạp, nhưng nếu chuyển sang một dạng ngôn ngữ khác, thì bài toán trở thành đơn giản.

Định nghĩa 1.1. Ta gọi dãy Catalan là dãy C_n được định nghĩa bởi công thức truy hồi

$$C_0 = 1,\ C_1 = 1;$$
$$C_n = C_0C_{n-1} + C_1C_{n-2} + \ldots + C_{n-1}C_0.$$

Định lý 1.1. Hàm sinh của dãy Catalan được xác định như sau:

$$F(x) = C_0 + C_1x + C_2x^2 \cdots + C_kx^k + \ldots = \frac{1-\sqrt{1-4x}}{2x}.$$

Chứng minh: Ta có $F(x) = \sum_{k=0}^{\infty} C_kx^k$. Khi đó

$$F(x) = C_0 + \sum_{k=1}^{\infty} C_kx^k = 1 + \sum_{k=1}^{\infty} C_kx^k$$

Mặt khác, $F^2(x) = \left(\sum_{k=0}^{\infty} C_kx^k\right)^2 = \sum_{k=0}^{\infty} a_kx^k$, với $a_k = \sum_{i=0}^{k} C_iC_{k-i} = C_{k+1}$. Suy ra

$F^2(x) = \sum_{k=0}^{\infty} C_{k+1}x^k = \frac{1}{x}\sum_{k=1}^{\infty} C_kx^k$, hay

$\sum_{k=1}^{\infty} C_kx^k = xF^2(x)$. Do đó $F(x) = 1 + xF^2(x)$.

Từ đó ta được $F(x) = \frac{1-\sqrt{1-4x}}{2x}$ hoặc

$$F(x) = \frac{1+\sqrt{1-4x}}{2x} \quad (*).$$

Do biểu thức (*) không khai triển được chuỗi lũy thừa tại $x = 0$ nên ta được $F(x) = \frac{1-\sqrt{1-4x}}{2x}$ là hàm sinh của dãy Catalan.

Hệ quả 1.2. Đẳng thức sau đúng: $C_n = \frac{1}{n+1}C_{2n}^n$.

Chứng minh: Ta khai triển hàm $F(x) = \frac{1-\sqrt{1-4x}}{2x}$. Theo định lý Newton mở rộng, ta có

$(1+x)^{\frac{-1}{2}} = \sum_{k=0}^{\infty} C_{\frac{-1}{2}}^k x^k$. Như vậy,

$$(1+x)^{\frac{-1}{2}} = \sum_{k=0}^{\infty} (-1)^k \frac{C_{2k}^k}{4^k} x^k.$$

Thay x bởi $-4x$ ta được $\frac{1}{\sqrt{1-4x}} = \sum_{k=0}^{\infty} C_{2k}^k x^k$.

Ta có $\sqrt{1-4x} = \frac{1}{\sqrt{1-4x}} - \frac{4x}{\sqrt{1-4x}}$.

Suy ra $\sqrt{1-4x} = \sum_{k=0}^{\infty} C_{2k}^k x^k - 4x\sum_{k=0}^{\infty} C_{2k}^k x^k$

$$=1+\sum_{k=1}^{\infty}C_{2k}^{k}x^{k}-4\sum_{k=0}^{\infty}C_{2k}^{k}x^{k+1}$$

$$=1+\sum_{k=1}^{\infty}C_{2k}^{k}x^{k}-4\sum_{k=1}^{\infty}C_{2k-2}^{k-1}x^{k}$$

$$=1+\sum_{k=1}^{\infty}(C_{2k}^{k}-4C_{2k-2}^{k-1})x^{k}.$$

Mặt khác, $C_{2k}^{k}-4C_{2k-2}^{k-1}=\dfrac{(2k)!}{(k!)^2}-4C_{2k-2}^{k-1}$

$$=\frac{(2k-2)!}{((k-1)!)^2}\cdot\frac{(2k-1)2k}{k^2}-4C_{2k-2}^{k-1}$$

$$=C_{2k-2}^{k-1}\left(\frac{(2k-1)2k}{k^2}-4\right)=-2\frac{C_{2k-2}^{k-1}}{k}.$$

Suy ra $\dfrac{1-\sqrt{1-4x}}{2x}=\sum_{k=1}^{\infty}\dfrac{C_{2k-2}^{k-1}}{k}x^{k-1}=\sum_{k=0}^{\infty}\dfrac{C_{2k}^{k}}{k+1}x^{k}$.

Vậy, công thức tổng quát của dãy Catalan là

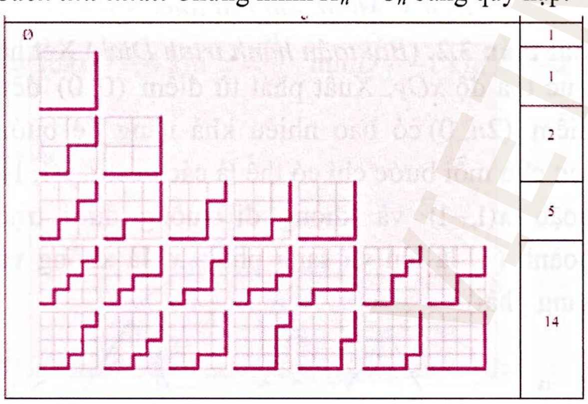
$$C_n=\frac{1}{n+1}C_{2n}^{n} \text{ với } n=0,1,2,\ldots$$

2. Các bài toán liên quan đến dãy Catalan

Bài toán 2.1. (*Bài toán bàn cờ*) Có bao nhiêu cách bước trên bàn cờ $n \times n$ từ ô phía dưới cùng bên trái đến ô trên cùng bên phải sao cho không bao giờ bước qua hẳn đường chéo chính, mỗi lần bước chỉ có thể lên 1 đơn vị hoặc sang phải 1 đơn vị, không bước lùi hay sang trái?

Lời giải. Kí hiệu H_n là số cách đi từ tọa độ (0, 0) đến tọa độ (n, n) theo đúng điều kiện của bài toán – không vượt qua đường chéo chính. Ta sẽ chứng minh bài này bằng hai cách.

Cách thứ nhất. Chứng minh $H_n = C_n$ bằng quy nạp.



Hình 1: Minh họa Bài toán 2.1

Ta có $H_0=1, H_1=1, H_2=2, H_3=5, H_4=14$.

Giả sử đẳng thức đúng đến n, tức là $H_n=H_0H_{n-1}+H_1H_{n-2}+\cdots+H_{n-1}H_0$. Ta phải chứng minh đẳng thức cũng đúng với $n+1$, tức là

$$H_{n+1}=H_0H_n+H_1H_{n-1}+\cdots+H_nH_0.$$

• *Trường hợp* 1. Khi di chuyển không bao giờ chạm vào đường chéo chính. Khi đó bước đầu tiên di chuyển phải là (0, 0) → (1, 0) và bước cuối cùng di chuyển phải là (n+1, n) → (n + 1, n + 1). Trong khoảng hai bước này, tình trạng chuyển động tương ứng với chuyển động của H_n. Như vậy trường hợp này có H_0H_n cách di chuyển.

• *Trường hợp* 2. Khi di chuyển có chạm vào đường chéo chính. Xét một chuyển động H_n như vậy. Kí hiệu (i, i) là vị trí lần đầu tiên khi chuyển động này chạm đường chéo chính. Rõ ràng, trong đoạn đầu từ (0,0) đến (i, i) chuyển động của ta nằm trong chuyển động H_i.

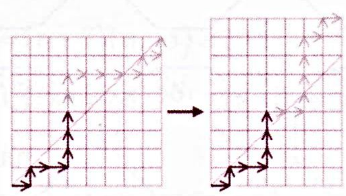
Từ (i, i) đến $(n+1, n+1)$ chuyển động của ta thuộc H_{n+1-i}. Do đó, chuyển động của H_n thuộc chuyển động của $H_{i-1}H_{n+1-i}$. Hiển nhiên chiều ngược lại cũng đúng, mỗi chuyển động thuộc $H_{i-1}H_{n+1-i}$ đều là một chuyển động của H_{n+1}. Kết hợp cả hai trường hợp ta có

$$H_{n+1}=H_0H_n+H_1H_{n-1}+\cdots+H_nH_0.$$

Theo nguyên lý quy nạp ta có $H_n=C_n$ với mọi n.

Cách thứ hai. Chứng minh H_n trực tiếp mà không sử dụng công thức truy hồi.

Không gây hiểu lầm nếu ta thay H_n bằng C_n.



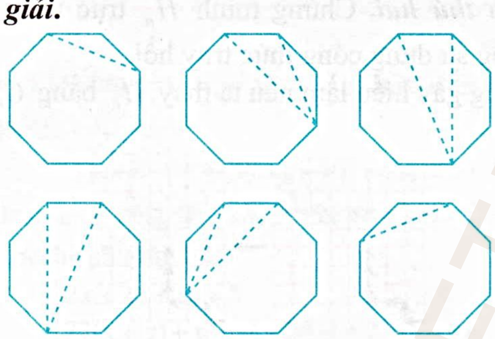
Hình 2: Minh họa Bài toán 2.1

Ta sẽ chứng minh có $\dfrac{1}{2n+1}C_{2n}^{n}$ cách đi từ góc cuối cùng bên trái lên góc trên cùng bên phải mà không bao giờ bước qua đường chéo chính (có thể chạm). Tổng số cách đi: C_{2n}^{n}. Ta sẽ tính cách đi phạm luật trước. Một cách đi gọi là phạm luật nếu bước qua đường chéo chính ít nhất một lần.

Ta xét một đường phạm luật tại lần đầu tiên. Khi đó quãng đường đi được đến nơi phạm luật sẽ phải có k bước sang phải và đúng $k+1$ bước lên trên. Như vậy để đi đến (n,n) thì cần $(n-k-1)$ bước đi lên và $(n-k)$ bước sang phải. Ta xét một phép biến hình sau: từ điểm phạm luật, ta lấy đối xứng qua đường chéo chính của quãng đường còn lại. Nếu đi sang phải thì đổi thành đi lên, và nếu là đi lên thì đổi thành sang phải. Vì phép biến hình này phân đường mới nhận được sẽ có $n-k$ bước đi lên và $n-k-1$ bước sang phải. Tổng cộng con đường mới sẽ có $k+1+(n-k)=n+1$ bước đi lên, và $k+(n-k-1)=n-1$ bước sang phải. Tức là điểm cuối cùng sẽ là $(n-1,n+1)$. Tất cả các con đường phạm luật sẽ có chung một điểm đến. Ta sẽ chỉ ra mỗi con đường từ $(0,0)$ đến $(n-1,n+1)$ đều tương ứng với một đường phạm luật. Cũng bằng phương pháp đối xứng qua đường chéo chính khi phạm luật đầu tiên, ta sẽ nhận được một lối đi phạm luật từ $(0,0)$ đến (n,n). Vậy, số đường phạm luật là C_{2n}^{n+1}. Suy ra số đường đúng luật là $C_n=C_{2n}^{n}-C_{2n}^{n+1}=\frac{1}{2n+1}C_{2n}^{n}$. Công thức được chứng minh.

Bài toán 2.2. (*Bài toán Euler*) Có bao nhiêu cách chia một đa giác lồi thành các tam giác mà không có cạnh nào cắt nhau?

Lời giải.



Hình 3: Minh họa Bài toán 2.2

Gọi giá trị phải tìm là E_n. Đặt $E_2=1$. Khi $n\geq 3$ ta có $E_3=1, E_4=2, E_5=5$. Đa giác đều n đỉnh được kí hiệu là $A_1A_2\ldots A_n$.

Xét một phân hoạch đa giác thành các tam giác. Xuất phát từ cạnh A_1A_n cố định, đỉnh thứ ba của tam giác có cạnh A_1A_n là A_k. Các đường chéo A_kA_1, A_kA_n chia đa giác $A_1A_2\ldots A_n$ thành ba phần: đa giác k đỉnh $A_1A_2\ldots A_k$, tam giác $A_1A_kA_n$ và đa giác $n-k+1$ đỉnh $A_kA_{k+1}\ldots A_n$. Số cách chia các đa giác còn lại sẽ là E_kE_{n-k+1} và phép tương ứng này với $A_1A_kA_n$ là song ánh. Cho k chạy từ 2 đến $n-1$ (với $n\geq 3$) nhận được: $E_n=\sum_{k=2}^{n-1}E_kE_{n-k+1}$.

Thay E_n bằng E_{n+2} vào công thức ta nhận được $E_{n+2}=\sum_{k=2}^{n+1}E_kE_{n-k+3}$. Thay E_{n+2} bằng E'_n vào công thức ta nhận được $E'_n=\sum_{k=2}^{n+1}E'_{k-2}E'_{n-k+1}$ hay $E'_n=\sum_{j=0}^{n-1}E'_jE'_{n-j-1}$. Công thức cuối chính là công thức Catalan. Do đó $E_n=C_{n-2}$.

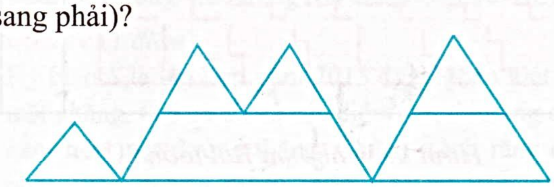
3. Các bài toán Catalan nổi tiếng khác

Bài toán 3.1. (*Bài toán dấu ngoặc*) Với n dấu ngoặc mở và n dấu ngoặc đóng (n bộ dấu ngoặc mở – đóng), có bao nhiêu cách sắp xếp các dấu ngoặc hợp lệ? (Hợp lệ ở đây được hiểu là kể từ trái sang phải, tại bất kì vị trí nào, thì số dấu ngoặc đóng đã sử dụng không vượt quá số ngoặc mở đã dùng và khi kết thúc thì đúng bằng nhau).

∅	**1**
()	**1**
(()), ()()	**2**
((())), (()()), (())(), ()(()), ()()()	**5**
(((()))), ((()())), ((())()), (()(())), (()()()), ((()))(), (()())(), (())(()), (())()(), ()((())), ()(()()), ()(())(), ()()(()), ()()()()	**14**

Hình 4: Minh họa Bài toán 3.1

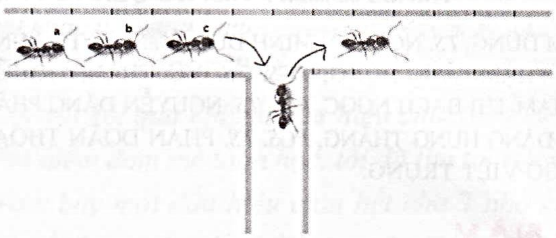
Bài toán 3.2. (*Bài toán hành trình Dick*) Xét hệ trục tọa độ xOy. Xuất phát từ điểm $(0, 0)$ đến điểm $(2n, 0)$ có bao nhiêu khả năng để bước sao cho mỗi bước chỉ có thể là các vectơ $y(1, 1)$ hoặc $x(1,-1)$ và không đi xuống dưới trục hoành (y là lên và sang phải, x là xuống và sang phải)?



Hình 5: Minh họa Bài toán 3.2

Bài toán 3.3. (*Bài toán phân vùng*) Có bao nhiêu bộ số x_i $(1\leq i\leq 2n)$, mỗi số x_i có giá trị là 1 hoặc -1 sao cho $x_1+x_2+\cdots+x_{2n}=0$ và tất cả các giá trị $x_1, x_1+x_2, x_1+x_2+x_3, \ldots, x_1+\cdots+x_{2n}$ đều không âm?

Bài toán 3.4. (*Bài toán đoàn quân kiến*) Có một đoàn quân kiến đang hành quân qua một con đường hầm chật hẹp. Không con nào có thể đổi chỗ cho nhau. Có một ngách nhỏ cũng chật hẹp như vậy, nếu con kiến nào muốn nghỉ thì có thể rẽ vào đó và các con khác lại tiếp tục đi. Nếu có nhiều con muốn nghỉ thì các con nghỉ trước lùi sâu vào ngách nhường chỗ cho con mới, theo thứ tự không được đổi chỗ. Khi ra thì trật tự ngược lại, con nào nghỉ sau thì ra trước. Hỏi có bao nhiêu cách ra khỏi đường hầm chật hẹp này?



Hình 6: Minh họa Bài toán 3.4

Bài toán 3.5. (*Bài toán mua vé xem phim*) Một lớp học có $2n$ học sinh đang đứng xếp hàng mua vé xem phim, giá vé là 10000 đồng/chiếc. Mỗi em học sinh chỉ có một trong hai tờ tiền là 10000 đồng và 20000 đồng.

Ban đầu quầy vé không có tiền. Có bao nhiêu cách mua vé để người bán vé luôn luôn trả lại được tiền thừa và công việc không bị gián đoạn?

Ta chứng minh các bài toán trên là tương đương:

– Bài toán Dấu ngoặc và bài toán Hành trình Dick: Với phép song ánh dấu ngoặc mở ứng với bước đi lên $y(1, 1)$ và dấu ngoặc đóng ứng với bước đi xuống $x(1, -1)$.

– Bài toán Hành trình Dick và bài toán Bàn cờ hoàn toàn là một khi ta xoay bàn cờ 45°.

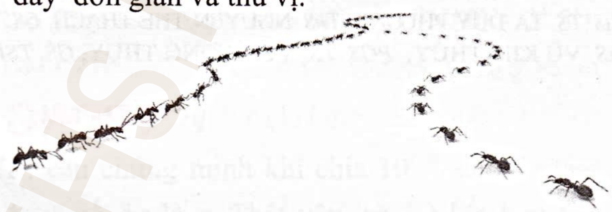
– Bài toán Phân vùng chính là bài toán Dấu ngoặc phát biểu dưới dạng đại số, thay ngoặc mở bằng +1 và ngoặc đóng bằng –1.

– Bài toán Đoàn quân kiến là bài toán Dấu ngoặc nếu với mỗi con kiến trước khi đến hẻm được phát một giấy kiểm tra (dấu ngoặc mở), sau khi nghỉ ngơi (hoặc tiếp tục đi luôn) qua hẻm thì thu hồi lại giấy thông hành (dấu ngoặc đóng).

– Tương tự như thế, bài toán Mua vé xem phim và bài toán Bàn cờ là tương đương khi trả tiền 10000 đồng tương ứng với bước sang phải 1 đơn vị và trả 20000 đồng tương ứng với bước lên trên 1 đơn vị.

Bài toán 3.6. (*Bài toán kiến hành quân*) Hai đoàn quân kiến vàng (m chiến sĩ) và kiến đen (n chiến sĩ) đang hành quân về điểm tập trung. Đến ngã ba thì hai đường hợp nhau thành một. Đèn đỏ dẫn đường không hoạt động. Có bao nhiêu cách hành quân qua ngã ba mà không chen lấn xô đẩy nhau?

Nếu thêm điều kiện số kiến đen được qua cửa luôn luôn không bé hơn số kiến vàng được qua cửa, ta cũng nhận được bài toán Catalan dưới đây đơn giản và thú vị.



Hình 7: Minh họa Bài toán 3.6

Bài toán mở. (*Bài toán Catalan tổng quát*) Có n đoàn quân kiến nhập làm một theo quy tắc số kiến của đoàn i luôn luôn không nhỏ hơn số kiến của đoàn k khi nhập hàng $(k \geq i)$. Có bao nhiêu cách thực hiện?

4. Một số bài tập

Bài toán 4.1. Có bao nhiêu hàm số $f: 1,2\ldots,n \to 1,2,\ldots n$ sao cho f là một hàm tăng và với mọi k ta có $f(k) \leq k$?

Bài toán 4.2. Có bao nhiêu cách bắt tay nhau của $2n$ người ngồi quanh một cái bàn tròn mà không có cặp nào bắt chéo tay với cặp nào?

Bài toán 4.3. Chứng minh rằng $C_n = C_{2n}^{n} - C_{2n}^{n+1}$.

Bài toán 4.4. Có 15 học sinh nam và 12 học sinh nữ cùng bước vào phòng khiêu vũ, hỏi có bao nhiêu cách vào? Biết rằng ở bất kì thời điểm nào, số học sinh nam đều không ít hơn số học sinh nữ. Tổng quát với m nam, n nữ và $m \geq n$ (trong bài chỉ phân biệt nam – nữ, không xét từng cá nhân cụ thể).