# Phân tích và thiết kế giải thuật

Khoa Khoa học máy tính

# Giới thiệu

- Khái niệm giải thuật/thuật toán (algorithm)
  - Thuật toán là một dãy xác định các thao tác cơ bản áp dụng trên dữ liệu vào nhằm đạt được giải pháp cho một vấn đề
  - Hai vấn đề
    - Tìm một phương pháp giải quyết vấn đề
      - Giải pháp cho $ax^2 + bx + c = 0$ : rỏ ràng và xác định
      - Giải pháp cho $ax^5 + bx^4 + cx^3 + dx^2 + ex + f = 0$ : không có giải pháp tổng quát
    - Tìm một giải pháp **hiệu quả**
  - Phân biệt giải thuật và chương trình
    - Chương trình là cài đặt thuật toán bằng một ngôn ngữ lập trình

# Giới thiệu

- Thuật toán
  - Thủ tục tính toán nhận tập các **dữ liệu vào** (input) và tạo các **dữ liệu ra** (output)
  - Thuật toán được gọi là **đúng đắn** (correct), nếu **thuật toán dừng** cho **kết quả đúng** với mọi dữ liệu vào

# Giới thiệu

- Thuật toán hằng ngày trong cuộc sống
  - Nấu cơm
    - ...
  - Gọi điện thoại
- Thuật toán trong toán học/tin học
  - Nhân hai ma trận
  - Tính tích phân
  - Giải hệ phương trình bậc nhất
  - ...

# Giới thiệu

- Các tính chất của thuật toán
  - Tính tổng quát
  - Tính hữu hạn
  - Tính không nhập nhằng
  - Tính hiệu quả

# Giới thiệu

- Tính tổng quát (1)
  - Thuật toán phải áp dụng cho tất cả các mẫu dữ liệu vào chứ không chỉ là một mẫu dữ liệu vào cụ thể
  - Ví dụ
    - Sắp xếp tăng dần một dãy các giá trị
    - Dữ liệu vào : 2 1 4 3 5
    - Kết quả : 1 2 3 4 5

# Giới thiệu

- Tính tổng quát (2)
  - Phương pháp
    - So sánh hai phần tử liên tiếp nhau từ trái qua phải, nếu không đúng thứ tự thì hoán đổi vị trí chúng

| | | | | | |
|---|---|---|---|---|---|
| Bước 1: | 2 ↔ | 1 | 4 | 3 | 5 |
| Bước 2: | 2 | 1 | 4 | 3 | 5 |
| Bước 3: | 1 | 2 | 4 ↔ | 3 | 5 |
| Bước 4: | 1 | 2 | 3 | 4 | 5 |

    - Dãy đã được sắp xếp
    - Thuật toán có tổng quát không ?
    - Không
      - Ví dụ dữ liệu vào: 2 1 5 4 3

# Giới thiệu

- Tính hữu hạn
  - Thuật toán phải dừng sau một số bước xác định
  - Ví dụ

```
nhập n
while (n ≠ 0)
    n = n – 2
endwhile
```

  - Thuật toán có dừng không ?

# Giới thiệu

- Tính không nhập nhằng
  - Các thao tác trong thuật toán phải được đặc tả chặt chẽ
    - Ở mỗi bước thực thi, bước tiếp theo sẽ được thực thi phải được xác định rõ ràng
  - Ví dụ

    Bước 1: x = 0
    Bước 2: tăng x lên 1 hoặc giảm x xuống 1
    Bước 3: <u>if</u> (x $\in$[-2,2]) <u>then</u>
    <u>goto</u> Bước 2

  - Thuật toán có nhập nhằng ?

# Giới thiệu

- Tính hiệu quả
  - Thuật toán phải sử dụng hiệu quả nguồn tài nguyên máy tính
    - Thời gian
    - Bộ nhớ
  - Ví dụ

```
gt1 (n)
begin
   p = 1
   for i from 1 to n do
       p = p * i;
   endfor
end
```

```
gt2 (n)
begin
   if (n = 0) return (1)
   else
      return (n*gt(n-1))
   endif
end
```

  - Thuật toán nào hiệu quả hơn ?

# Giới thiệu

- Đặc tả thuật toán (1)
  - Có nhiều cách đặc tả thuật toán
    - Không hình thức
      - Ngôn ngữ tự nhiên
    - Nữa hình thức
      - Kêt hợp ngôn ngữ tự nhiên và các kí hiệu toán học
        - Sơ đồ khối, ngôn ngữ giả, …
    - Hình thức
      - Các kí hiệu toán học
        - Ngôn ngữ Z, ngôn ngữ B, …

# Giới thiệu

- Đặc tả thuật toán (2)
  - Chọn phương pháp đặc tả nữa hình thức
    - Ngôn ngữ giả tựa C/Pascal kết hợp ngôn ngữ tự nhiên

```
if (điều kiện) then
    …
else
    …
endif
```

```
while () do
    …
endwhile
```

```
for … from … to … do
    …
endfor
```

# Giới thiệu

- Ví dụ 1
  - Có 15 hộp đinh có chiều dài khác nhau. Cần đặt chúng vào trong 15 ngăn kéo. Chúng ta cần phải thực hiện thế nào để có thể lấy được loại đinh chúng ta cần một cách nhanh nhất?

# Giới thiệu

- Ví dụ 1
  - Giải pháp thô
    - Chúng ta xếp các hộp đinh vào các ngăn kéo theo thứ tự bất kỳ. Sau đó, mỗi khi cần lấy đinh chúng ta mở các ngăn kéo theo thứ tự từ trái qua phải.
    - Trường hợp xấu nhất: phải mở 15 ngăn kéo
    - Trường hợp trung bình: giả sử xác xuất lấy các loại đinh ngang nhau, phải mở 8 ngăn kéo

$$\frac{1}{15}\sum_{1}^{15} i = \frac{1}{15} * \frac{15*16}{2} = 8$$

    - Nếu những loại đinh trong các ngăn kéo cuối cùng luôn được sử dụng, cần nhiều lần mở các ngăn kéo

# Giới thiệu

- Ví dụ 1
  - Giải pháp Las Vegas
    - Chúng ta chọn ngăn kéo để mở một cách ngẫu nhiên
      - Nếu may mắn, chúng ta có loại đinh cần lấy
      - Nếu không, chúng ta loại bỏ ngăn kéo vừa mở, thực hiện mở ngăn kéo khác một cách ngẫu nhiên

$$\frac{1}{15}*1+\frac{14}{15}*\frac{1}{14}*2+\frac{14}{15}*\frac{13}{14}*\frac{1}{13}*3+\ldots=\frac{1}{15}\sum_{1}^{15} i=8$$

# Giới thiệu

- Ví dụ 1
  - Giải pháp tiền xử lý
    - Chúng ta sắp xếp các hộp đinh theo thứ tự chiều dài của chúng
    - Xếp vào các ngăn kéo theo thứ tự trên
    - Nếu lại mở các ngăn kéo từ trái sang phải thì sẽ không thu được lợi ích gì
    - Mở các ngăn kéo theo phương pháp tìm kiếm nhị phân
    - Trường hợp xấu nhất: 4 lần mở ngăn kéo
    - Trường hợp trung bình: 3.26

$$\frac{1}{15}*1+\frac{2}{15}*2+\frac{4}{15}*3+\frac{8}{15}*4=3.26$$

    - Giải pháp chỉ có ý nghĩa khi chúng ta thường xuyên mở các ngăn kéo lấy đinh

# Giới thiệu

- Ví dụ 2
  - Thuật toán nhân hai số nguyên
  - Thuật toán nhân Bắc Mỹ

```
      5 6 7
    1 2 3 4
-----------
    2 2 6 8
  1 7 0 1
1 1 3 4
5 6 7
-----------
6 9 9 6 7 8
```

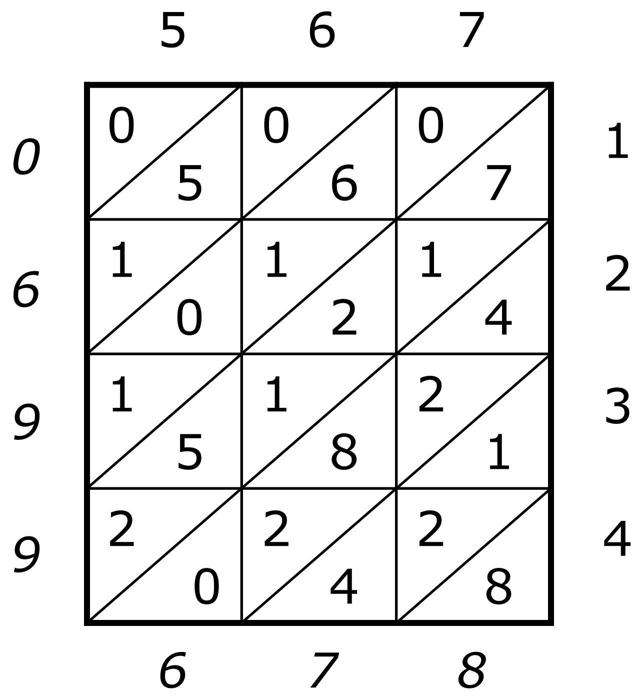
# Giới thiệu

- Ví dụ 2
  - Thuật toán nhân Anh

```
          5 6 7
        1 2 3 4
---------------
    5 6 7
    1 1 3 4
      1 7 0 1
        2 2 6 8
---------------
    6 9 9 6 7 8
```

# Giới thiệu

- Ví dụ 2
  - Thuật toán nhân Ả-rập




  - So sánh độ hiệu quả các thuật toán ?

# Giới thiệu

- Ví dụ 2
  - Đếm số phép toán nhân và cộng được thực hiện
    - Thuật toán nhân Bắc Mỹ và thuận toán nhân Anh đều sử dụng 12 phép nhân và 15 phép cộng
    - Thuật toán Ả-rập sử dụng 12 phép nhân và 20 phép cộng
    - Số phép toán phụ thuộc vào số chữ số của mỗi số nguyên
      - Số phép nhân bằng m*n với m và n lần lượt là số chữ số của mỗi số nguyên

# Giới thiệu

- Ví dụ 2
  - Thuật toán nhân Nga

| | |
|---|---|
| 567 | 1234 |
| 283 | 2468 |
| 141 | 4936 |
| *70* | *9872* |
| 35 | 19744 |
| 17 | 39488 |
| *8* | *78976* |
| *4* | *157952* |
| *2* | *315904* |
| 1 | 631808 |
| | 699678 |

**Ưu điểm**: không cần ghi nhớ các kết quả nhân trung gian, chỉ cần thực hiện phép cộng và chia cho 2

# Giới thiệu

- Ví dụ 2
  - Thuật toán nhân « chia để trị »
    - Số chữ số của hai số nguyên phải bằng nhau và phải bằng luỹ thừa 2
      - Nếu số chữ số của hai số nguyên khác nhau thì thêm vào các số 0 ở bên trái
    - Ý tưởng: chúng ta thay phép nhân của hai số nguyên có n chữ số bởi 4 phép nhân của hai số nguyên có n/2 chữ số. Tiếp tục, thay phép nhân của hai số nguyên có n/2 chữ số bởi 4 phép nhân của hai số nguyên có n/4 chữ số, …
    - Cần nhân hai số nguyên: a và b. Gọi $a_L$ và $a_R$ là hai nữa trái và phải của số a, tương tự $b_L$ và $b_R$ là hai nữa trái và phải của b

# Giới thiệu

- Ví dụ 2
  - Thuật toán nhân « chia để trị »

$$
\begin{array}{rccc}
 & & a_L & a_R \\
 & * & b_L & b_R \\
\hline
 & & a_L b_R & a_R b_R \\
+ & a_L b_L & a_R b_L & \\
\hline
\end{array}
$$

$$a_L b_L + (a_L b_R + a_R b_L) + a_R b_R$$

$$
\begin{aligned}
a*b &= (a_L*10^{n/2} + a_R)*(b_L*10^{n/2} + b_R) \\
&= a_L*b_L*10^{n} + a_L*b_R*10^{n/2} + a_R*b_L*10^{n/2} + a_R*b_R \\
&= a_L*b_L*10^{n} + (a_L*b_R + a_R*b_L)10^{n/2} + a_R*b_R
\end{aligned}
$$

Để nhân a và b bởi các nữa của a và b, cần sử dụng 4 phép nhân

# Giới thiệu

- Ví dụ 2
  - Thuật toán nhân « chia để trị »
    - Giảm số phép nhân a và b bởi các nữa của a và b từ 4 xuống còn 3

      Đặt: $p = a_L*b_L$
      $q = a_R*b_R$
      $r = (a_L + a_R)*(b_L + a_R)$
      Khi đó:
      $a*b = a_L*b_L*10^n + (a_L*b_R + a_R*b_L)*10^{n/2} + a_R*b_R$
      $= p*10^n + (r – p - q)*10^{n/2} + q$

    - Số phép nhân sử dụng bởi thuật toán này được đánh giá là $n*m^{0.59}$ với $n \geq m$
    - Khi nhân hai số nguyên khá lớn, thuật toán này được đánh giá nhanh hơn các thuật toán trước

# Giới thiệu

- Ví dụ 2
  - Thuật toán nhân « chia để trị »

0567 * 1234

| nhân | số bước dịch | kết quả |
|---|---|---|
| 05 12 | 4 | 60 |
| 05 34 | 2 | 170 |
| 67 12 | 2 | 804 |
| 67 34 | 0 | 2278 |
| | | 699678 |

05 * 34

| nhân | số bước dịch | kết quả |
|---|---|---|
| 0 3 | 2 | 0 |
| 0 4 | 1 | 0 |
| 5 3 | 1 | 15 |
| 5 4 | 0 | 20 |
| | | 170 |

# Giới thiệu

- Ví dụ 3: tính $x^n$
  - Vào: một số nguyên x và một số nguyên không âm n
  - Ra: $y = x^n$
  - Thuật toán đơn giản

```
if n = 0 then
        y = 1
else
        y = x
endif
for i from 1 to n do
        y = y * x
endfor
```

# Giới thiệu

- Ví dụ 3: tính $x^n$
  - Thuật toán nhị phân
    - Giảm số phép toán
    - Ý tưởng: sử dụng các kết quả trung gian
    - Nguyên tắc: nếu n/2 = 0 thì $x^n = x^{n/2} * x^{n/2}$ nếu không thì $x^n = x^{n-1} * x$
    - Chẳng hạn
      - $x^{35} = x^{34} * x$
      - $x^{34} = x^{17} * x^{17}$
      - $x^{17} = x^{16} * x$
      - $x^{16} = x^8 * x^8$
      - $x^8 = x^4 * x^4$
      - $x^4 = x^2 * x^2$
      - $x^2 = x * x$

# Giới thiệu

- Ví dụ 3: tính $x^n$
  - Thuật toán nhị phân
    1. Biểu diễn n bởi dãy nhị phân
    2. Thay mỗi bít nhị phân

       « 1 » bởi các kí hiệu SX

       « 0 » bởi các kí hiệu X
    3. Xoá cặp kí hiệu SX bên trái nhất
    4. Kết quả: cách tính $x^n$ trong đó

       S nghĩa là bình phương

       X nghĩa là nhân với x

       Bắt đầu từ x

# Giới thiệu

- Ví dụ 3: tính $x^n$
  - Thuật toán nhị phân

```
Biểu diễn n bởi dãy nhị phân: e_{k-1}e_{k-2}...e_0
if e_{k-1} = 0 then
        y = 1 // khi n = 0
else
        y = x
endif
for i from k-2 to 0 do
        y = y * y
        if e_i = 1 then
                y = y * x
        endif
endfor
```

# Giới thiệu

- Ví dụ 3: tính $x^n$
  - Thuật toán nhị phân
    - Ví dụ: $n = 35 = 100011_{(2)}$

| $e_i$ | y | số phép nhân |
|---|---|---|
| 1 | x | 0 |
| 0 | $x^2$ | 1 |
| 0 | $x^4$ | 2 |
| 0 | $x^8$ | 3 |
| 1 | $x^{17}$ | 5 |
| 1 | $x^{35}$ | 7 |

# Giới thiệu

- Ví dụ 3: tính $x^n$
  - Thuật toán nhị phân: giải thích
    - Biểu diễn nhị phân của n: $n = \sum_{i=0}^{p} A_i 2^i$
    - Giả sử chúng ta đang trong quá trình tính $x^n$. Gọi j là vị trí bít bên cuối cùng (trái nhất) biểu diễn n, $y_j$ là kết quả cuối cùng có được. Ban đầu: j = p, $y_p = x = x^{Ap}$
    - Có hai khả năng của $A_{j-1}$
      - $A_{j-1} = 1$. Thay thế $A_{j-1}$ bởi SX. Vậy $y_{j-1} = y_j^2 * x$
      - $A_{j-1} = 0$. Thay thế $A_{j-1}$ bởi S. Vậy $y_{j-1} = y_j^2$

      Cả hai trường hợp: $y_{j-1} = y_j^2 * x^{Aj-1}$
    - Vậy: $y_{p-1} = y_p^2 * x^{Ap-1} = (x^{Ap})^2 * x^{Ap-1} = x^{2Ap + Ap-1}$
    - Bằng truy hồi:

$$y_1 = x^{\sum_{i=0}^{p} A_i 2^i} = x^n$$

# Giới thiệu

- Ví dụ 3: tính $x^n$
  - Thuật toán nhị phân: có giải pháp tốt hơn ?
    - Chẳng hạn tính $x^{15}$
      1. n = 15 = 1111
      2. SX SX SX SX
      3. SX SX SX
      4. Bắt đầu bởi x, chúng ta có: $x^2$, $x^3$, $x^6$, $x^7$, $x^{14}$, $x^{15}$

      Vậy, chúng ta sử dụng 6 phép nhân

      Tuy nhiên, chúng ta có thể tính $x^{15}$ bởi:

      $x^2$, $x^3$, $x^6$, $x^{12}$, $x^{15} = x^{12} * x^3$

      Nghĩa là chỉ cần sử dụng 5 phép nhân

# Giới thiệu

- Ví dụ 3: tính $x^n$
  - Phương pháp ước số
    - $x^n = x$ nếu $n = 1$
    - $x^n = x^{n-1} * x$ nếu n là số nguyên tố
    - $x^n = (x^r)^s$ nếu $n = r * s$ với r là ước số nguyên tố nhỏ nhất của n và $s > 1$
  - Ví dụ: $n = 15$
    - $15 = 3 * 5$, khi đó $x^{15} = (x^3)^5$. Áp dụng thuật toán để tính $x^3$ và $c^5$ với $c = x^3$.
    - $x^3 = x^2 * x$, áp dụng thuật toán tính $x^2$
    - $x^2 = x * x$
    - $c^5 = c^4 * c$, áp dụng thuật toán tính $c^4$
    - $c^4 = (c^2)^2$
    - 2 phép nhân tính $c^4$, 3 phép nhân tính $c^5$, 2 phép nhân tính $x^3$. Vậy 5 phép nhân tính $x^{15}$

# Giới thiệu

- Phân tích và đánh giá các thuật toán
  - Dựa vào các tính chất của thuật toán
  - Chứng minh sự đúng đắn
  - Đánh giá độ phức tạp