

Nhập Môn Công Nghệ Thông Tin

Bài 05

CHƯƠNG TRÌNH MÁY TÍNH & THUẬT TOÁN

Giảng viên: ...

Email: ...



Khoa Công Nghệ Thông Tin
Trường Đại Học Khoa Học Tự Nhiên

Nội Dung

- Chương trình máy tính
- Chương trình dịch
- Thuật toán
- Thông tin – dữ liệu – tri thức

CHƯƠNG TRÌNH MÁY TÍNH

Khái niệm

- Chương trình lưu trữ (stored program)
 - Có thể được mã hóa dưới dạng chuỗi bits và được ***lưu trong bộ nhớ***
 - Là ***chuỗi các câu lệnh*** thực hiện 1 thuật toán
 - Từ đó, CPU có thể trích ra các câu lệnh và đem đi thực thi
 - Là một dạng đặc biệt của dữ liệu

Câu lệnh là gì?

- Khi yêu cầu người khác thực hiện điều gì đó, ta dùng *ngôn ngữ của con người*
 - Rườm rà và mơ hồ đối với máy tính để có thể giải mã (decode) và thực thi (execute)
- *Ngôn ngữ máy*
 - Ngắn gọn (concise) : dễ dàng giải mã
 - Rõ ràng (precise) : 1 câu lệnh chỉ được thực hiện bởi 1 cách

Thuật ngữ liên quan

- Mã lệnh (machine instruction)
 - Là lệnh (command hay instruction) được mã hóa dưới dạng chuỗi bits có thể nhận biết bởi CPU
- Ngôn ngữ máy (machine language)
 - Là một tập lệnh được nhận biết bởi 1 máy

Ngôn ngữ máy

- Reduced Instruction Set Computing (RISC)
 - Ít lệnh, đơn giản, hiệu quả và nhanh
 - Ví dụ:
 - PowerPC của Apple/IBM/Motorola
 - ARM
- Complex Instruction Set Computing (CISC)
 - Nhiều lệnh, thuận lợi, mạnh mẽ
 - Ví dụ: Intel

Loại mã lệnh

- Data transfer
 - Sao chép dữ liệu từ 1 địa điểm tới 1 địa điểm khác
- Arithmetic/Logic
 - Sử dụng các chuỗi bits có sẵn để tính toán ra chuỗi bits mới
- Control
 - Chỉ đạo thực hiện chương trình

Ví dụ : cộng

Step 1. Get one of the values to be added from memory and place it in a register.

Step 2. Get the other value to be added from memory and place it in another register.

Step 3. Activate the addition circuitry with the registers used in Steps 1 and 2 as inputs and another register designated to hold the result.

Step 4. Store the result in memory.

Step 5. Stop.

Nguồn: Computer Science - An Overview, 12e

Ví dụ : chia

Step 1. LOAD a register with a value from memory.

Step 2. LOAD another register with another value from memory.

Step 3. If this second value is zero,
JUMP to Step 6.

Step 4. Divide the contents of the first register by the second register and leave the result in a third register.

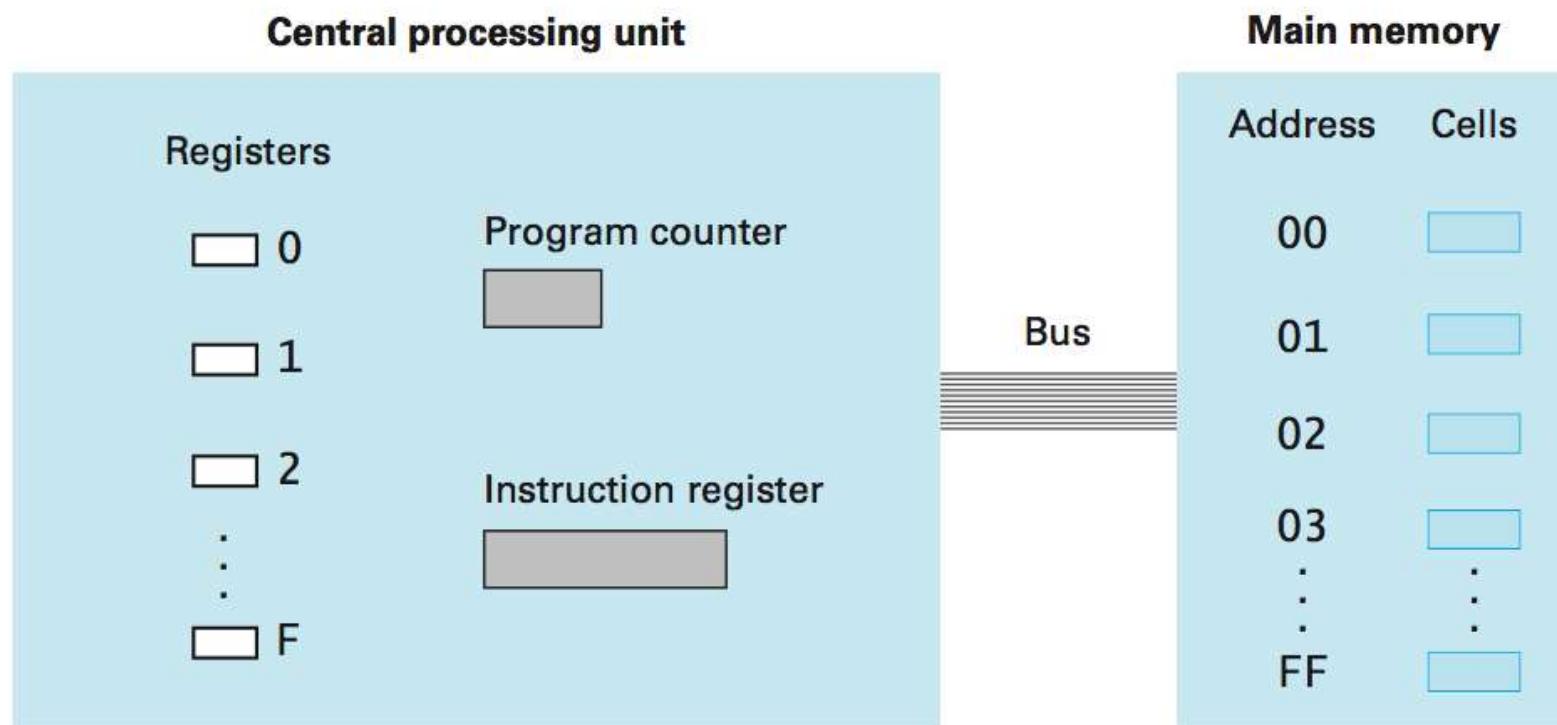
Step 5. STORE the contents of the third register in memory.

Step 6. STOP.

Nguồn: Computer Science - An Overview, 12e



Ví dụ : kiến trúc máy



Nguồn: Computer Science - An Overview, 12e

Thành phần của mã lệnh

- Chuỗi bits của mã lệnh gồm 2 phần
 - Op-code field
 - Chỉ định thao tác nào được thực thi
 - Operand field
 - Cung cấp thêm thông tin chi tiết về thao tác
 - Dữ liệu và địa chỉ
 - Số lượng operands tùy thuộc vào op-code

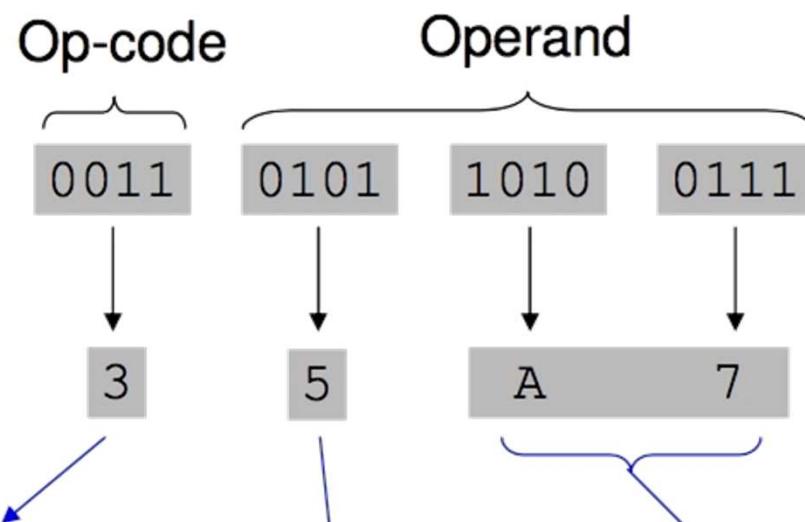
Ví dụ

- Câu lệnh yêu cầu CPU
 - Cộng hai ô nhớ data2 và data6
 - Sau đó lưu kết quả và ô nhớ data7
- Yêu cầu này được mã hóa bởi 16 bits



Source: Chun-Jen Tsai, ics12, National Chiao Tung University

Ví dụ cụ thể



Op-code 3 means to store the contents of a register in a memory cell

This part of the operand identifies the register whose contents are to be stored

Actual bit pattern (16 bits)

Hexadecimal form (4 bits)

This part of the operand identifies the address of the memory cell that is to receive data

Ví dụ cộng ở mã lệnh

Encoded instructions	Translation
156C	Load register 5 with the bit pattern found in the memory cell at address 6C.
166D	Load register 6 with the bit pattern found in the memory cell at address 6D.
5056	Add the contents of register 5 and 6 as though they were two's complement representation and leave the result in register 0.
306E	Store the contents of register 0 in the memory cell at address 6E.
C000	Halt.

Nguồn: Computer Science - An Overview, 12e



Tập mã lệnh đơn giản

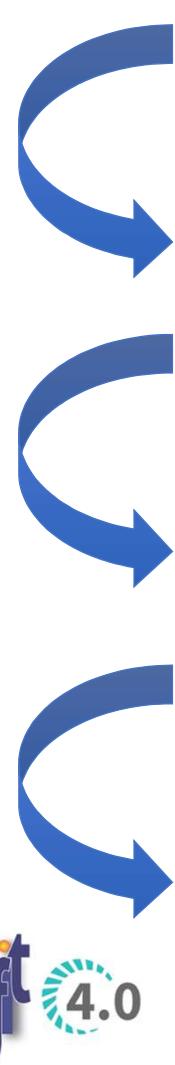
Op-code	Operand	Description
1	RXY	LOAD reg. R from cell XY
2	RXY	LOAD reg. R with XY
3	RXY	STORE reg. R at XY
4	0RS	MOVE R to S
5	RST	ADD S and T into R (2's comp.)
6	RST	ADD S and T into R (floating pt.)
7	RST	OR S and T into R
8	RST	AND S and T into R
9	RST	XOR S and T into R
A	R0X	ROTATE reg. R X times
B	RXY	JUMP to XY if R = reg. 0
C	0	HALT

Ví dụ

- 14A3
- 20A3
- 35B1
- 40A4
- 5726
- 634E
- 7CB4
- 8045
- 95F3
- A403
- B43C
- C000

TRÌNH BIÊN DỊCH – TRÌNH THÔNG DỊCH

Giới thiệu



High level language

Assembly Language

Machine Language

HARDWARE



mov ecx, ebx

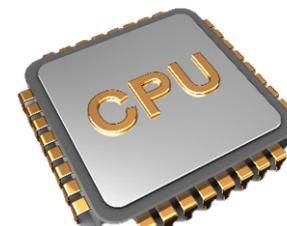
mov esp, edx

mov edx, r9d

100101011001

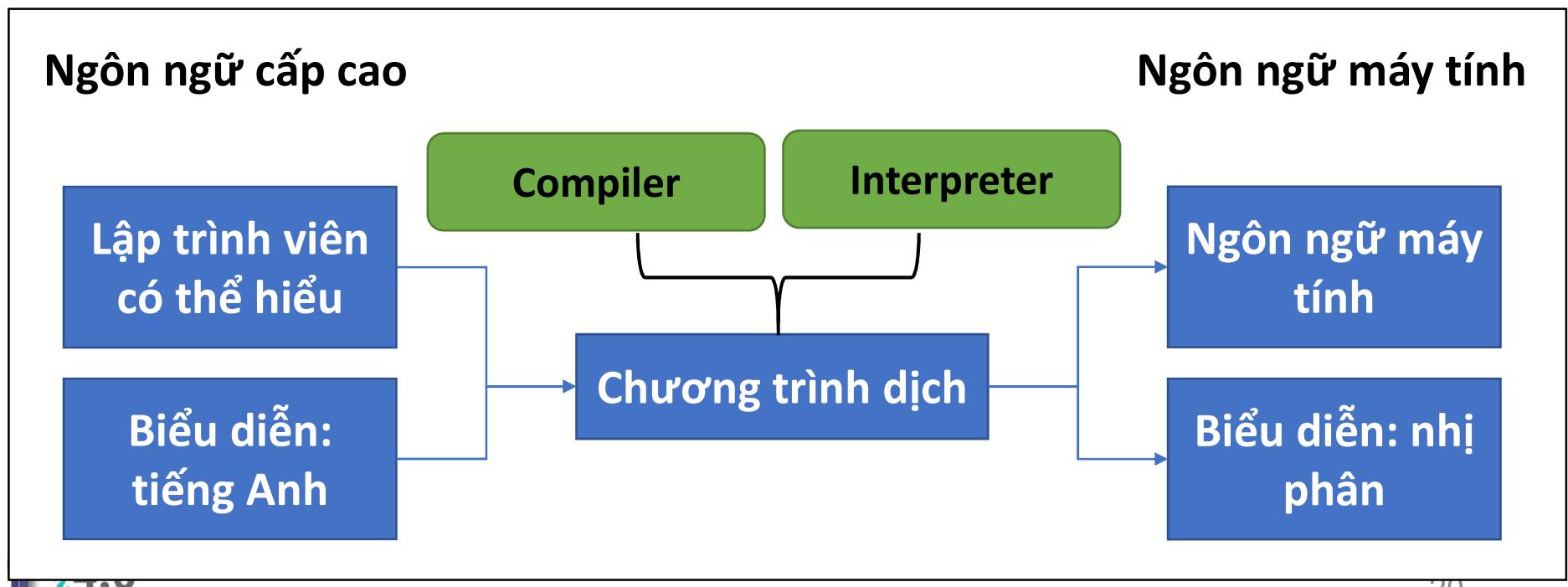
01001111011

111010101101

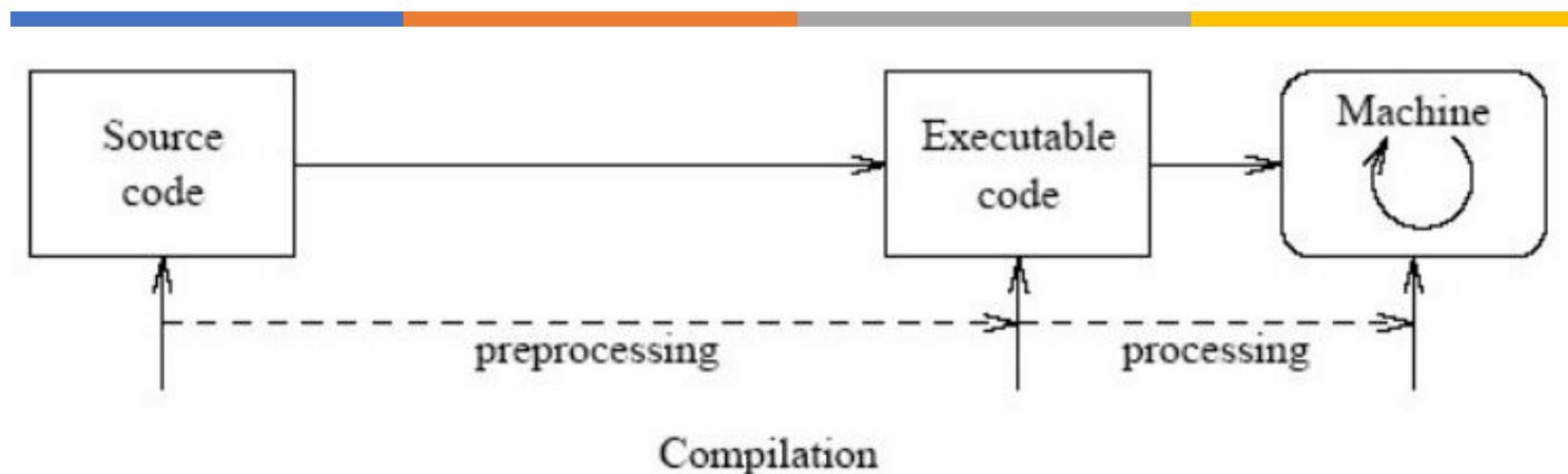


Mục đích

- Trình biên dịch (Compiler) và trình thông dịch (Interpreter) có cùng mục đích là dịch chỉ thị từ ngôn ngữ cấp cao thành ngôn ngữ máy

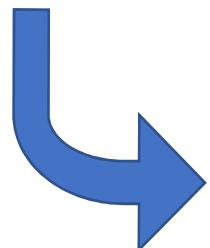


Cấu trúc chương trình dịch



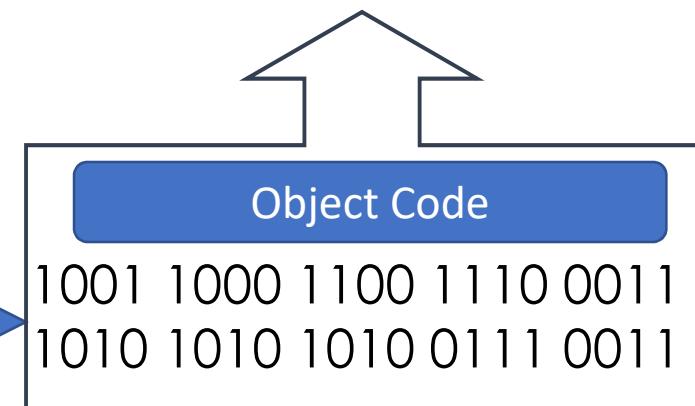
Original Source Code

Pctoutput.isVisible = true
Pctoutput.print "Hello"

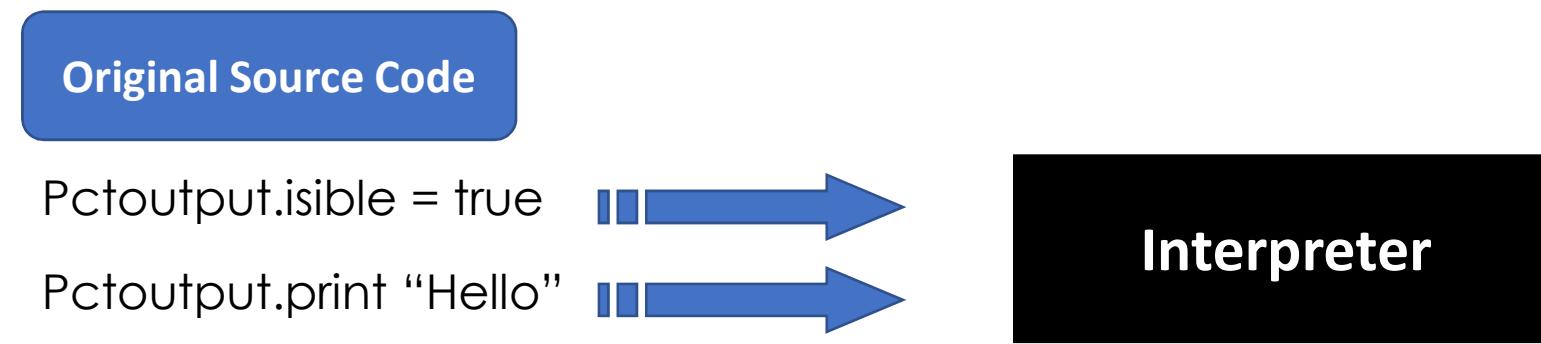
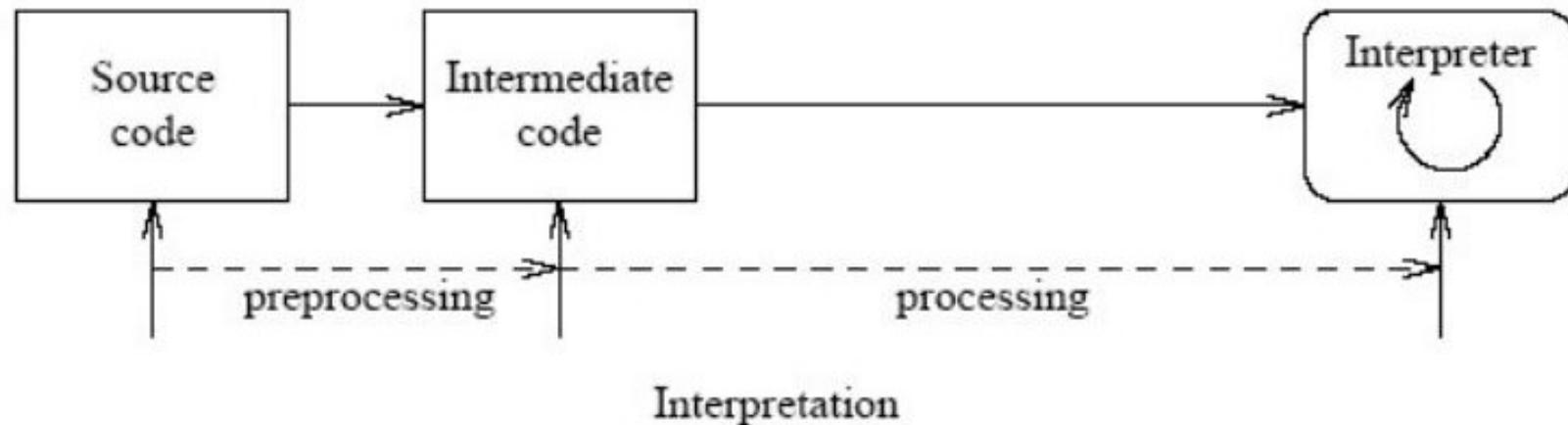


Compiler

Can be saved and
ran later



Cấu trúc chương trình dịch



Khác biệt

Tiêu chí	Compiler	Interpreter
Input	Toàn bộ chương trình	Từng dòng lệnh
Output	Phát sinh mã trung gian	Không phát sinh mã trung gian
Speed	Nhanh	Chậm
Memory	Yêu cầu nhiều vùng nhớ	Dùng ít vùng nhớ hơn vì không cần tạo mã trung gian
Working mechanism	Biên dịch trước khi thực thi	Biên dịch và thực thi cùng lúc
Errors	Hiển thị tất cả lỗi sau khi biên dịch	Hiển thị lỗi sau mỗi dòng lệnh
Error detection	Khó	Tương đối dễ
Programming language	C, C++, C#, Scala,...	PHP, Perl, Python, Ruby,...

THUẬT TOÁN

Giới thiệu

- CPU tuân theo chu kỳ máy tính

As long as the halt instruction has not been executed continue to execute the following steps:

- a. Fetch an instruction.
- b. Decode the instruction.
- c. Execute the instruction.

Thuật toán đơn giản



Source: Computer Science - An Overview, 12e

Định nghĩa

- Thuật toán là một *tập có thứ tự* các bước *rõ ràng, có thể thực thi* mà xác định được *quá trình dừng*
- Thuật toán/giải thuật là một tập hợp hữu hạn của các chỉ thị hay phương cách được định nghĩa cho việc hoàn tất một số việc từ một trạng thái ban đầu cho trước dẫn đến kết quả mong muốn
- Một bài toán có thể được giải quyết bởi các thuật toán khác nhau

Ví dụ

- Thuật toán để giải phương trình bậc nhất $P(x)$:
 $ax + b = c$ (với a, b, c là các số thực) có thể
thực hiện qua một số bước sau

Nếu $a = 0$

$b = c$ thì $P(x)$ có nghiệm bất kì

$b \neq c$ thì $P(x)$ vô nghiệm

Nếu $a \neq 0$

$P(x)$ có duy nhất một nghiệm $x = (c - b) / a$

Tầm quan trọng của thuật toán

- Theo Boolos và Jeffrey

- Không có con người nào có thể viết đủ nhanh, đủ dài, đủ nhỏ để liệt kê tất cả các thành phần của một tập rất lớn gần như vô hạn mà chỉ bằng cách lần lượt viết ra tên của chúng theo một số quy ước
- Tuy nhiên, con người có thể đưa ra cách thức để xác định phần tử thứ n bất kì
- Từ đó, cách thức này sẽ được thực hiện bởi các máy điện toán

Câu hỏi về thuật toán

- Các nhà phát triển thuật toán thường tự đặt 4 câu hỏi phản biện (critical) khi họ đánh giá các thuật toán
 - Có phải thuật toán giải quyết bài toán đã được nêu ra?
 - Có phải thuật toán rõ ràng, rành mạch?
 - Thuật toán có đưa ra một kết xuất?
 - Thuật toán có kết thúc trong một khoảng thời gian hợp lý?

PHÁT TRIỂN THUẬT TOÁN

Quá trình phát triển thuật toán

- Xác định đầu vào
- Xác định tiến trình thực hiện
- Xác định đầu ra
- Phát triển lược đồ HIPO
- Xác định các thành phần (module) liên quan

Quá trình phát triển thuật toán

- Xác định đầu vào
 - Thuật toán cần dữ liệu gì ?
 - Như thế nào để có dữ liệu đó ?
 - Định dạng dữ liệu như thế nào ?
- Xác định tiến trình thực hiện
- Xác định đầu ra
- Phát triển lược đồ HIPO
- Xác định các thành phần (module) liên quan

Quá trình phát triển thuật toán

- Xác định đầu vào
- Xác định tiến trình thực hiện
 - Làm cách nào để thao tác với dữ liệu để sinh ra những kết quả có ý nghĩa?
- Xác định đầu ra
- Phát triển lược đồ HIPO
- Xác định các thành phần (module) liên quan

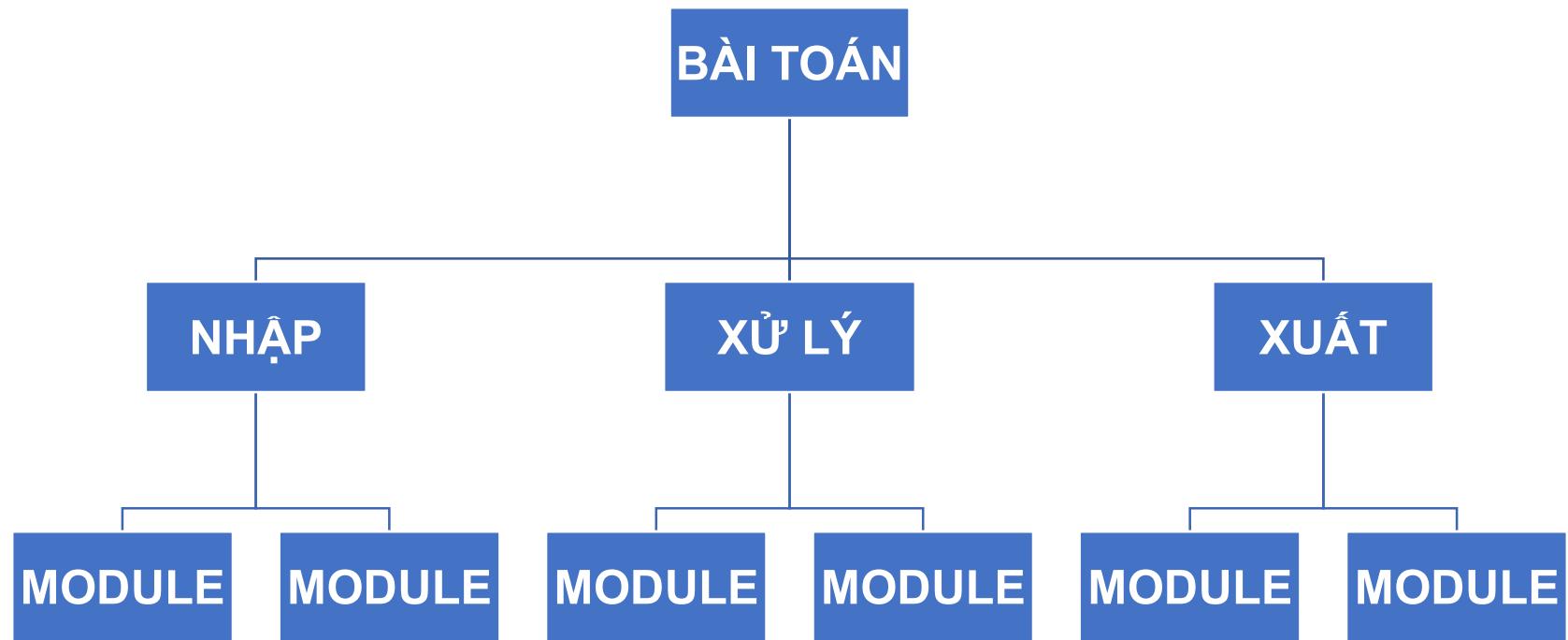
Quá trình phát triển thuật toán

- Xác định đầu vào
- Xác định tiến trình thực hiện
- Xác định đầu ra
 - Dữ liệu nào cần được trả ra?
 - Định dạng dữ liệu trả ra?
- Phát triển lược đồ HIPO
- Xác định các thành phần (module) liên quan

Phát triển lược đồ HIPO

- HIPO (Hierarchy of Input-Processes-Outputs) là một kĩ thuật phục vụ cho việc lên kế hoạch và ghi tài liệu cho thuật toán
- HIPO là một biểu đồ phân tầng thể hiện cấu trúc điều khiển và một bộ nhập-xử lý-xuất để mô tả dữ liệu đến, dữ liệu xuất từ đâu và những xử lý được thực thi bởi các module trên lược đồ phân tầng này

Phát triển lược đồ HIPO



Xác định module liên quan

- Như thế nào để tách những bài toán lớn thành những mảnh nhỏ hơn và có thể quản lý được?
- Các module cần dữ liệu đầu vào nào?
- Những xử lý cần được thực hiện trong mỗi module?
- Dữ liệu kết xuất của từng module?

BIỂU DIỄN THUẬT TOÁN

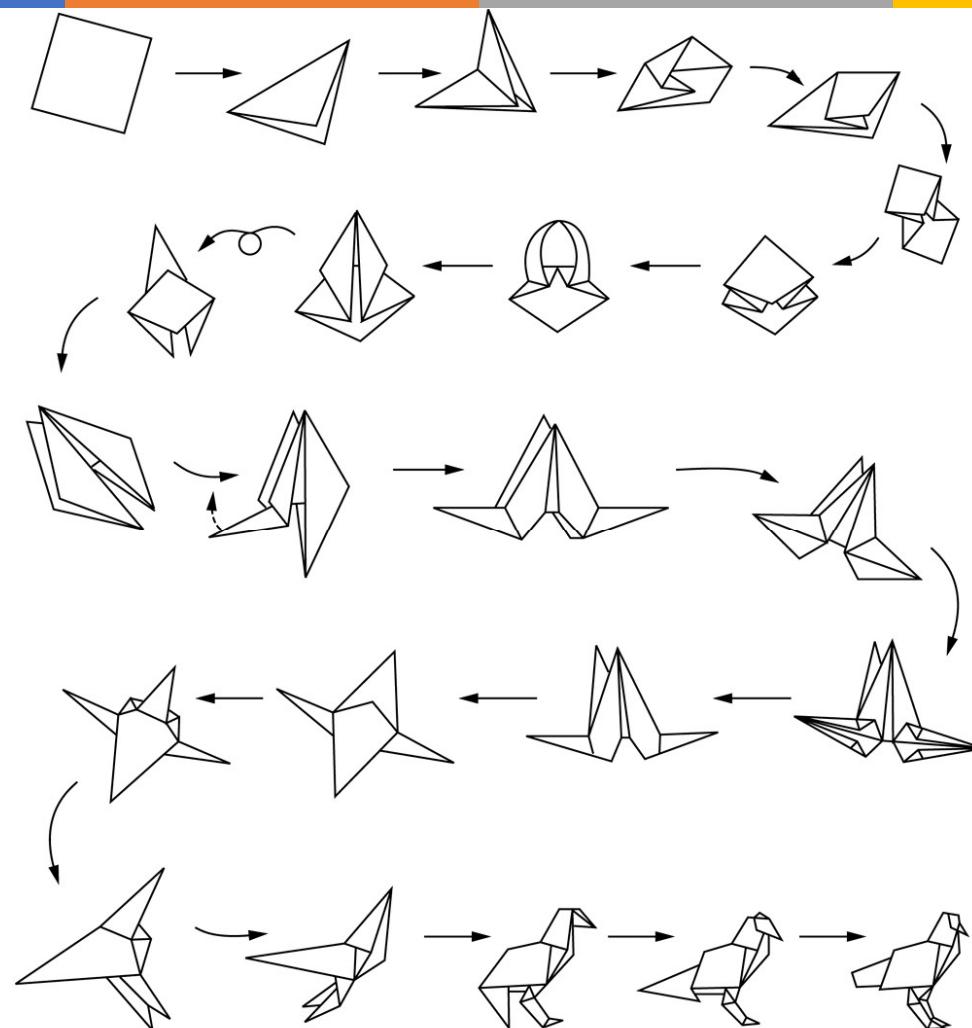
Biểu diễn thuật toán

- Cần 1 *ngôn ngữ* để biểu diễn thuật toán
 - Được định nghĩa chính xác
 - Cung cấp đầy đủ thông tin
- Xem xét
 - Ngôn ngữ tự nhiên: hiểu nhầm
 - Ngôn ngữ hình ảnh: mức độ chi tiết

Biểu diễn thuật toán

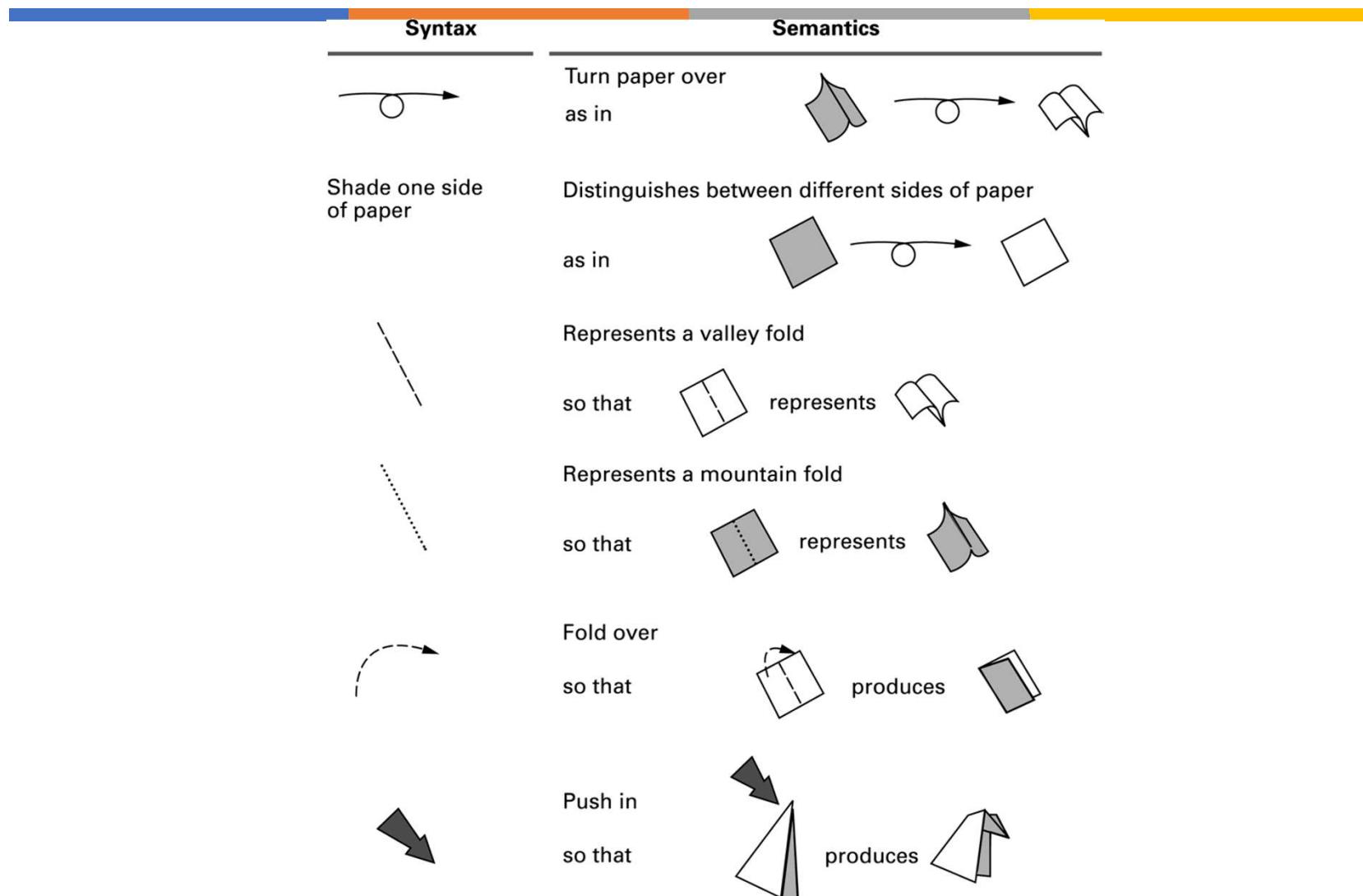
- Giải pháp
 - Nguyên thủy (primitives)
 - Cú pháp (syntax) - Ký hiệu
 - Ngữ nghĩa (semantic) - Ý nghĩa
 - Mã giả (pseudocode)
 - Hệ thống ký hiệu trực giác biểu diễn các ý tưởng thuật toán

Ví dụ



Source: Computer Science - An Overview, 12e

Ví dụ



Source: Computer Science - An Overview, 12e

Mã giả

- Phép gán

name = expression

- Ví dụ

RemainingFunds =
CheckingBalance + SavingsBalance

Mã giả

- Phép chọn có điều kiện

`if (condition):
 activity`

- Ví dụ

`if (sales have decreased):
 lower the price by 5%`

Mã giả

- Phép chọn có điều kiện

```
if (condition):  
    activity  
else:  
    activity
```

- Ví dụ

```
if (year is leap year):  
    daily total = total / 366  
else:  
    daily total = total / 365
```

Mã giả

- Thực hiện vòng lặp

`while (condition):
 body`

- Ví dụ

`while (tickets remain to be
sold):
 sell a ticket`

Mã giả

- Các điều kiện lồng nhau

```
if (not raining):  
    if (temperature == hot):  
        go swimming  
    else:  
        play golf  
else:  
    watch television
```

Mã giả

- Định nghĩa hàm

```
def name():
```

- Ví dụ

```
def ProcessLoan():
```

- Gọi hàm

```
if (...):  
    ProcessLoan()  
else:  
    RejectApplication()
```

Mã giả

- Thủ tục

```
def Greetings():
    Count = 3
    while (Count > 0):
        print('Hello')
        Count = Count - 1
```

Mã giả

- Tham số

```
def Sort(List):
```

- Ví dụ

```
Sort(the membership list)
```

```
Sort(the wedding guest list)
```

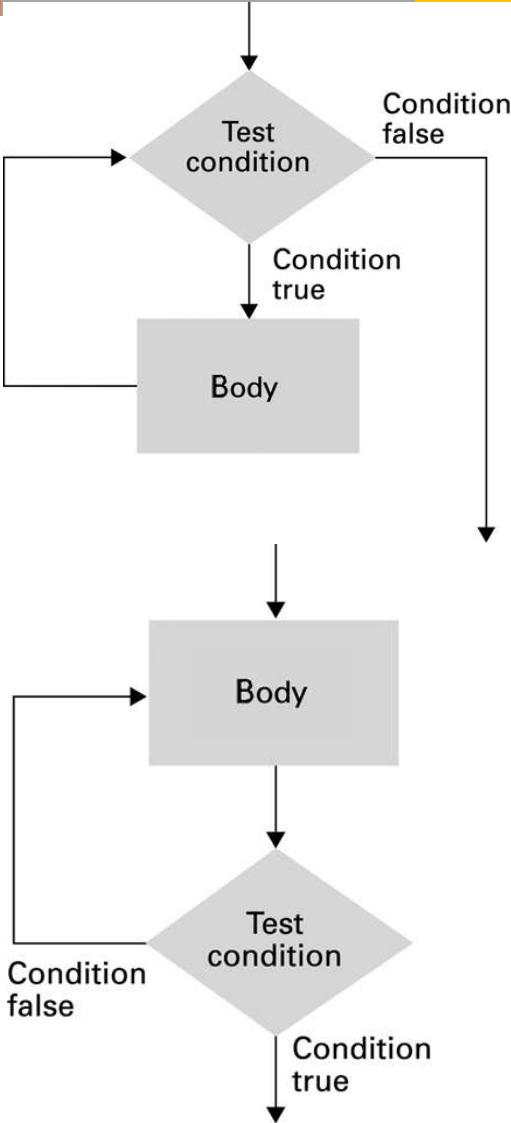
Biểu diễn thuật toán

- Giải pháp
 - Lược đồ
 - Bảng điều khiển
 - Ngôn ngữ lập trình

Lược đồ

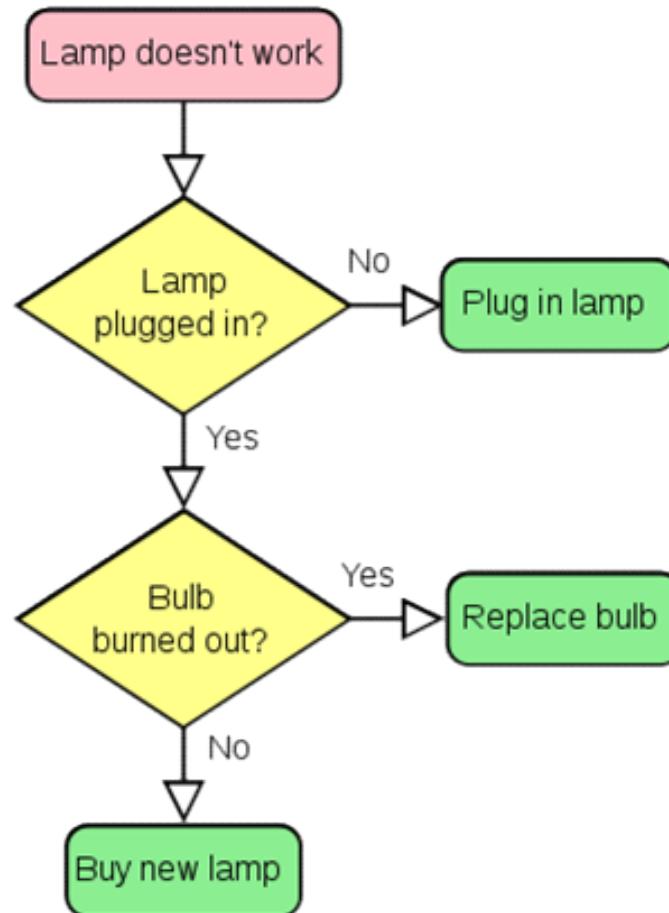
- Vòng lặp
while (*condition*)
body

repeat:
body
until(*condition*)

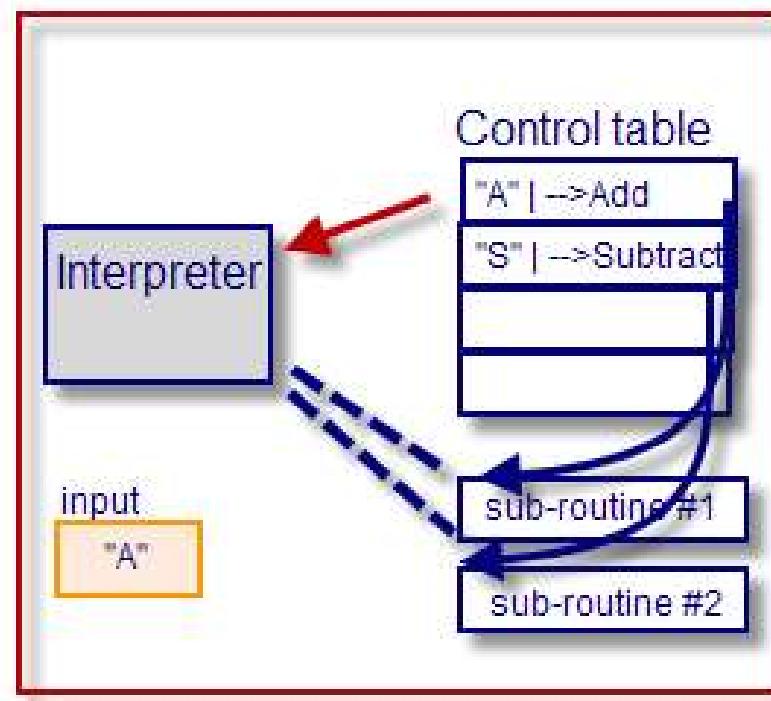


Source: Computer Science - An Overview, 12e

Lược đồ



Bảng điều khiển



ĐÁNH GIÁ THUẬT TOÁN

Đánh giá thuật toán

- Thuật toán được đánh giá dựa trên khối lượng tài nguyên (thời gian và bộ nhớ) cần để thực thi nó
 - Độ phức tạp về mặt không gian
 - Độ phức tạp về mặt thời gian

Tại sao cần đánh giá

- Việc sử dụng vô ý một thuật toán không hiệu quả có thể ảnh hưởng đến hiệu năng hệ thống
- Trong các ứng dụng thời gian thực, một thuật toán chạy quá lâu có thể làm cho kết quả của nó đã lỗi thời hoặc vô dụng
- Một thuật toán không hiệu quả cũng có thể tiêu tốn một khối lượng tính toán hay vùng nhớ một cách không kinh tế để chạy

Phân loại thuật toán

- Tiêu chí phân loại
 - Theo cách thực thi
 - Tuần tự, song song, ...
 - Theo phương pháp thiết kế
 - Vét cạn, chia để trị, ...
 - Theo lĩnh vực nghiên cứu
 - Tìm kiếm, sắp xếp, ...
 - Theo độ phức tạp
 - Khối lượng thời gian cần để hoàn thành so với kích thước dữ liệu nhập

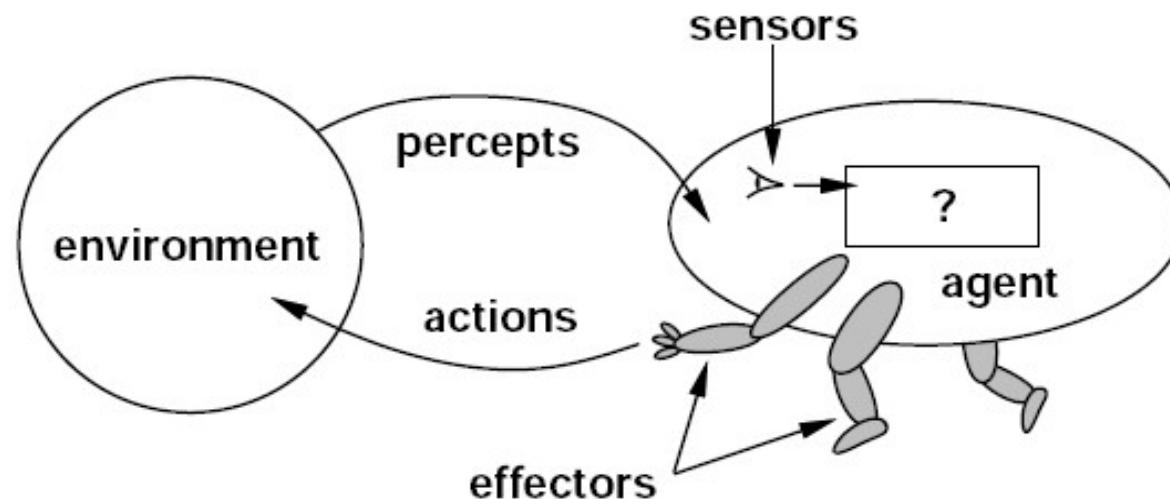
TRÍ TUỆ NHÂN TẠO

Trí tuệ nhân tạo là gì

- Trí tuệ nhân tạo là lĩnh vực khoa học máy tính
 - Tìm cách chế tạo máy có thể thực hiện các nhiệm vụ phức tạp mà không cần sự can thiệp của con người.
 - Máy móc có thể nhận thức và suy luận
- Trí tuệ nhân tạo liên quan đến nhiều ngành như tâm lý học, thần kinh học, toán học, ngôn ngữ học, và kỹ thuật điện và cơ khí.

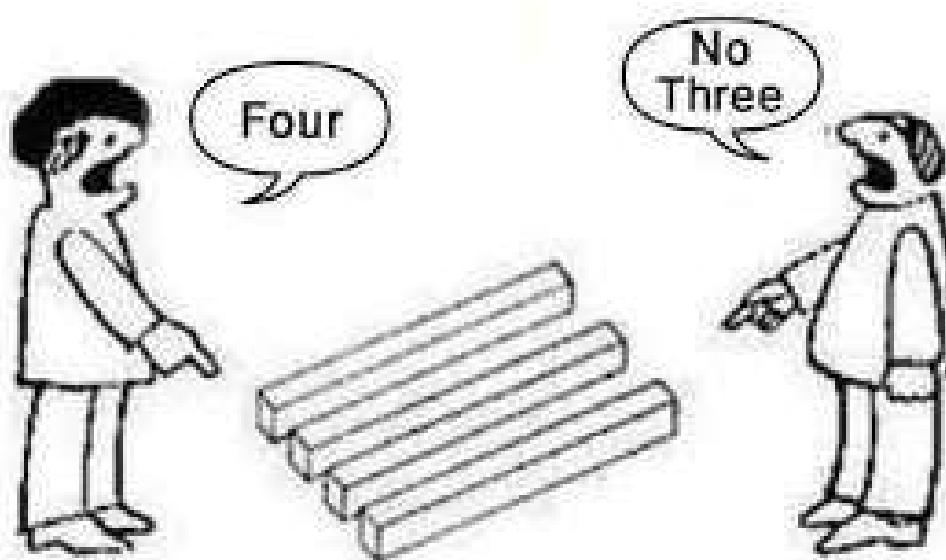
Tác nhân thông minh

- Tác nhân: Một thiết bị phản ứng lại với các kích thích từ môi trường xung quanh
 - Cảm biến (micrô, máy ảnh, cảm biến..)
 - Thiết bị truyền động (bánh xe, chân, cánh...)
- Phần lớn các nghiên cứu về trí tuệ nhân tạo là xây dựng các tác nhân hoạt động thông minh



Nhận thức

- Để phản ứng thông minh với đầu vào, một tác nhân phải có khả năng hiểu đầu vào đó (nhận thức).
- Nhận thức hình ảnh
- Nhận thức ngôn ngữ.

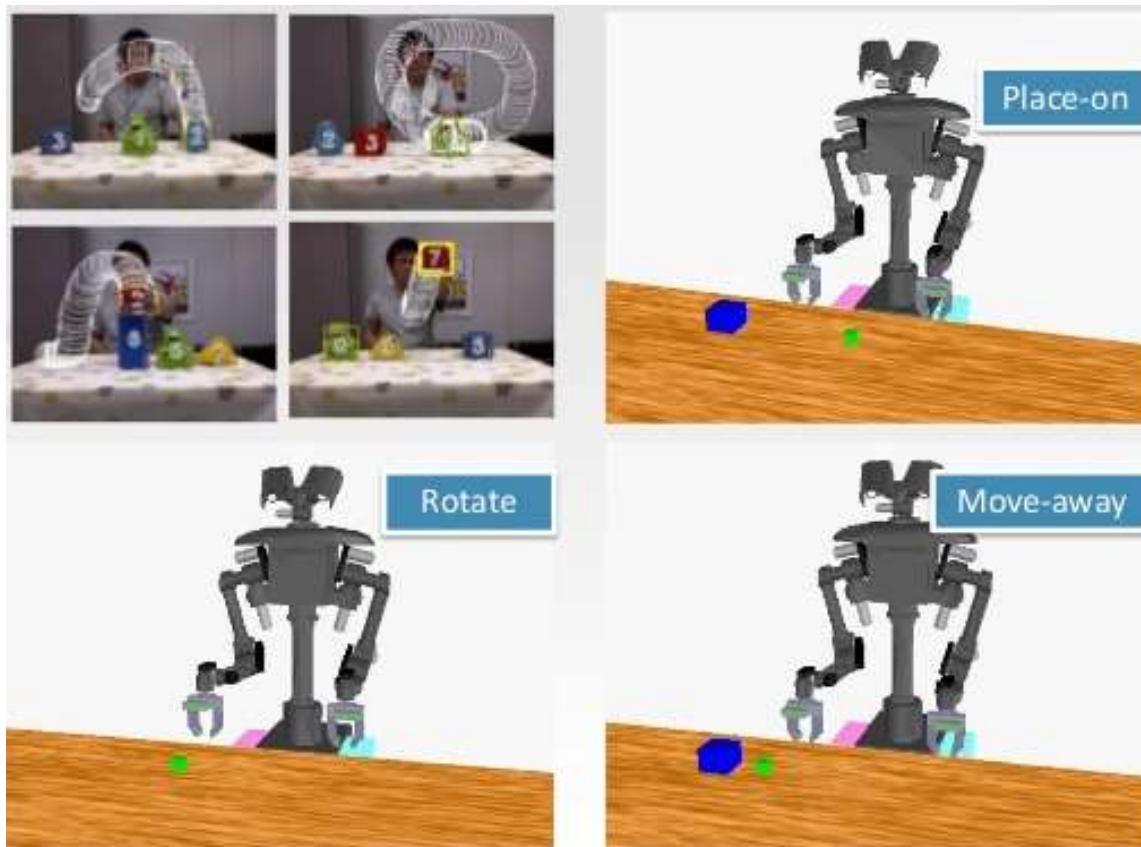


Suy luận

- Suy luận là hành động rút ra kết luận từ một số tiền đề bằng cách sử dụng một phương pháp nhất định.
- Suy luận là một quá trình suy nghĩ.
- Suy luận là lập luận hợp lý.
- Suy luận là tìm ra những gì nó cần biết từ những gì nó đã biết

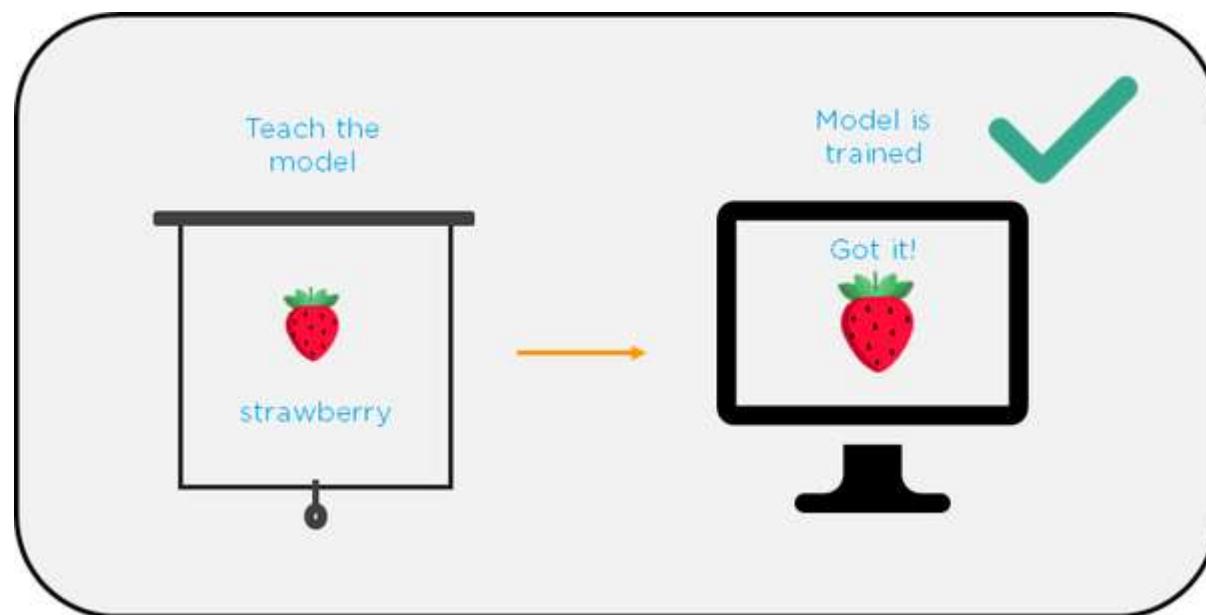
Học (learning)

- Sự bắt chước: trực tiếp thể hiện các bước trong một nhiệm vụ và máy tính chỉ cần ghi lại các bước.



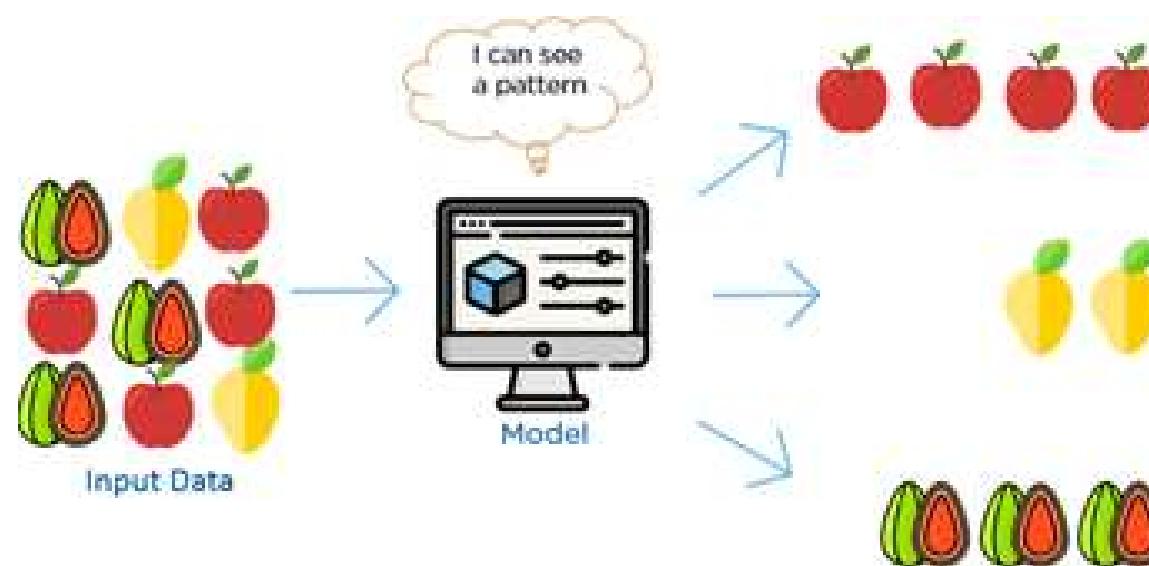
Học (learning)

- Học có giám sát: học một hàm ánh xạ đầu vào thành đầu ra dựa trên các cặp mẫu đầu vào-đầu ra.



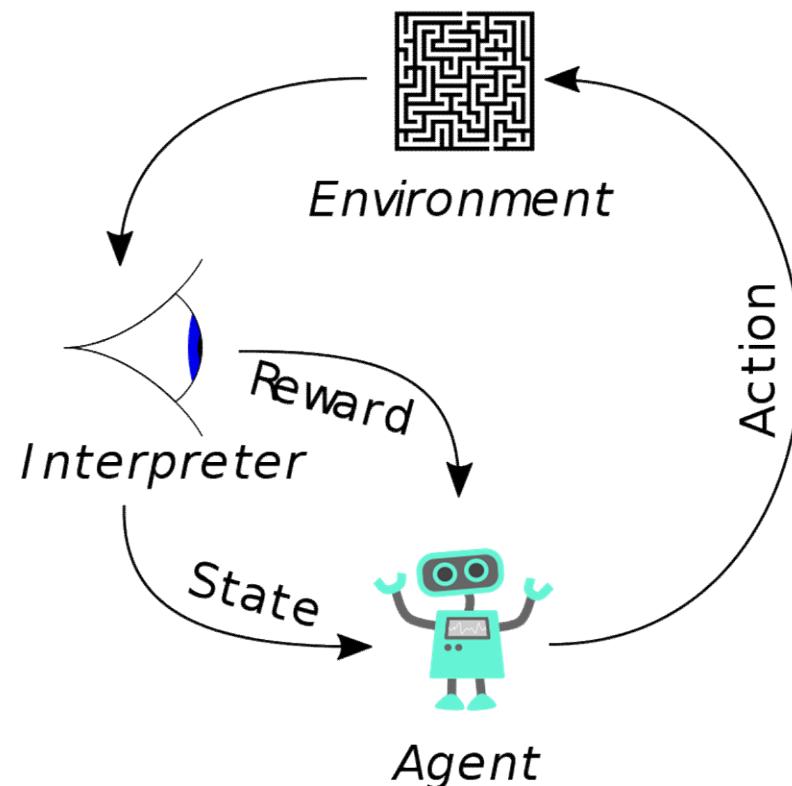
Học (learning)

- Học tập không giám sát: học từ dữ liệu chưa được dán nhãn.
- Việc học tập không giám sát xác định điểm tương đồng trong dữ liệu và phản ứng dựa trên sự hiện diện hay vắng mặt của những điểm tương đồng như vậy trong mỗi phần dữ liệu mới.



Học (learning)

- Học tăng cường: cách thực hiện các hành động trong một môi trường để tối đa hóa một tích lũy.



NGÀNH KHOA HỌC MÁY TÍNH

Nguồn nhân lực

- Tổng số GV: 14 cán bộ công tác trong nước
 - 2 PGS.TS
 - 4 Tiến sĩ
 - 8 Thạc sĩ
- **Trưởng bộ môn:** PGS.TS. Lê Hoài Bắc
- **Phó Trưởng bộ môn:** ThS. Lê Ngọc Thành

Nguồn nhân lực

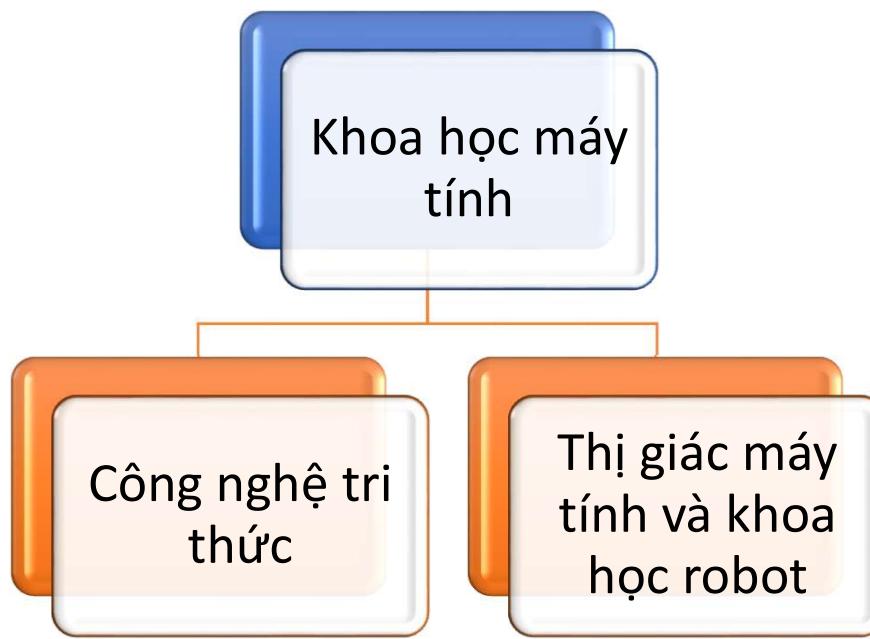


Các khái niệm

- Khoa học máy tính là ngành nghiên cứu các cơ sở lý thuyết về thông tin và tính toán cùng sự thực hiện và ứng dụng của chúng trong các hệ thống máy tính.
- Khoa học máy tính có nhiều phân ngành nhỏ, một số phân nhánh nhấn mạnh vào việc tính toán các kết quả cụ thể (vd đồ họa máy tính), trong khi các chi nhánh khác lại liên hệ đến tính chất của những vấn đề có thể giải quyết được dùng phương pháp máy tính, (vd như Lý thuyết độ phức tạp tính toán). Còn lại những chi nhánh khác thì tập trung vào những khó khăn trong việc thực thi những phương pháp dùng để tính toán (vd ngành Lý thuyết ngôn ngữ lập trình)

Các khái niệm

- Hiện trong khoa, ngành Khoa Học Máy Tính được chia làm 2 phân ngành hẹp hơn:
 - Ngành Công Nghệ Tri Thức
 - Ngành Thị Giác Máy Tính và Khoa Học Robot

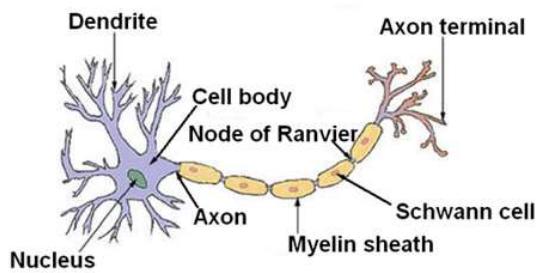


Mục tiêu đào tạo

- Cung cấp cho sinh viên các kiến thức trong lĩnh vực tính toán mềm, khai thác dữ liệu, sinh trắc học, máy học và nhận dạng mẫu, tính toán song song, ẩn dữ liệu.
- Kỹ năng quan tâm và chia sẻ (care & share)
- Khả năng tự tìm hiểu, nghiên cứu các công nghệ, phương pháp, quy trình mới trong lĩnh vực khoa học máy tính

Hướng nghiên cứu ngành

- ✓ Mạng thần kinh nhân tạo
- ✓ Lập trình tiến hóa
- ✓ Dự đoán luật
- ✓ Thuật giải di truyền
- ✓



An ứ
liệu



Tính toán song song

Chuyên
Ngành Khoa
Học Máy
Tính

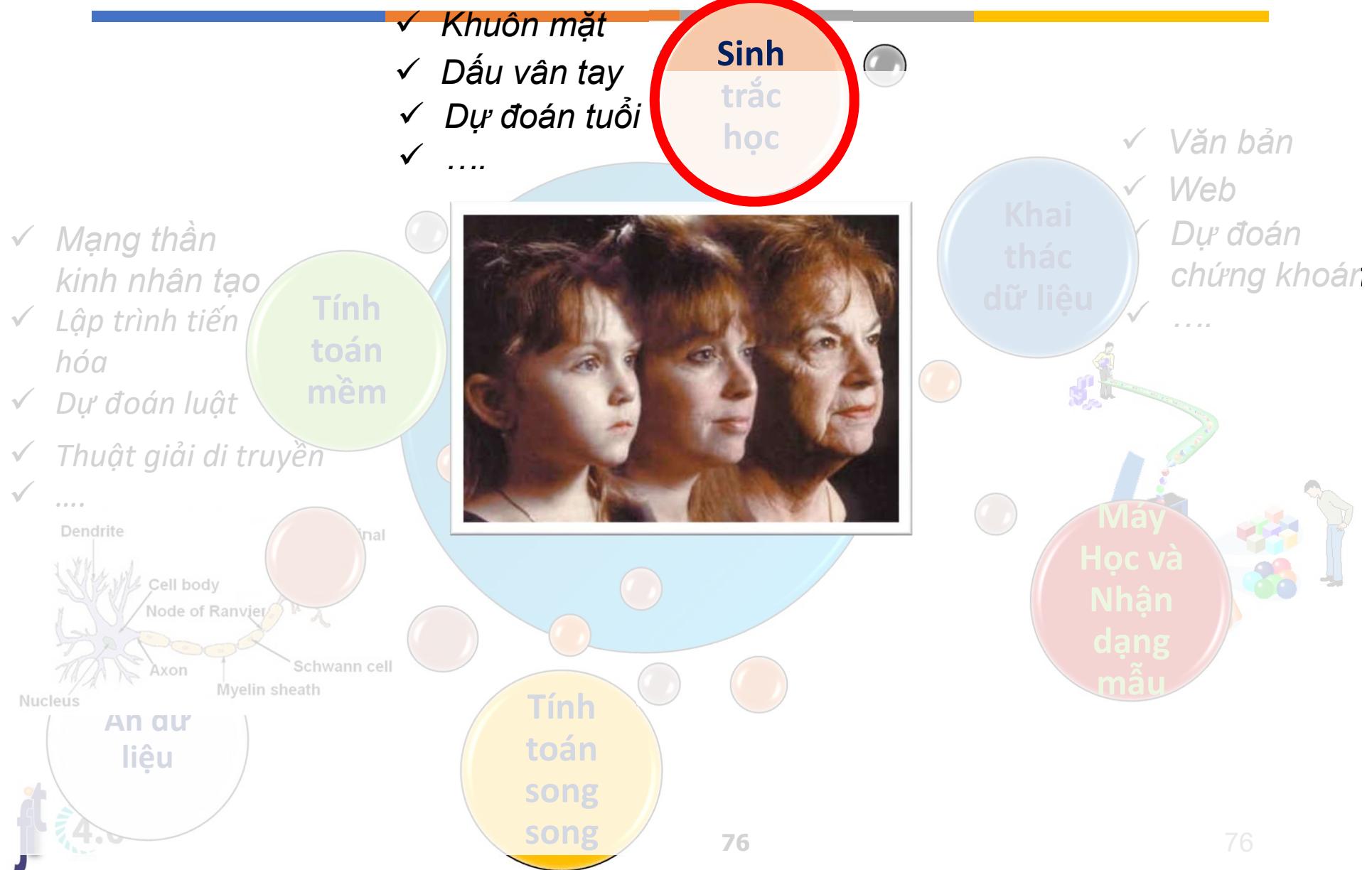
Sinh
trắc
học

Khai
thác
dữ liệu

Hướng nghiên cứu ngành



Hướng nghiên cứu ngành



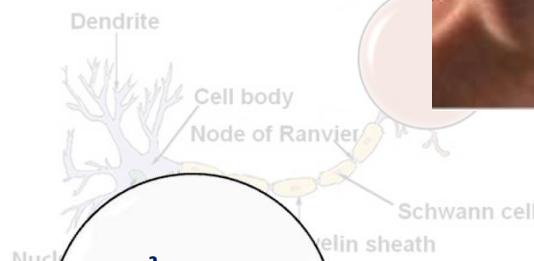
Hướng nghiên cứu ngành

✓ Khuôn mặt

✓ Đầu vân tay

Sinh

- ✓ Mạng thần kinh nhân tạo
- ✓ Lập trình tiến hóa
- ✓ Dự đoán luật
- ✓ Thuật giải di truyền
- ✓



Ẩn dữ liệu



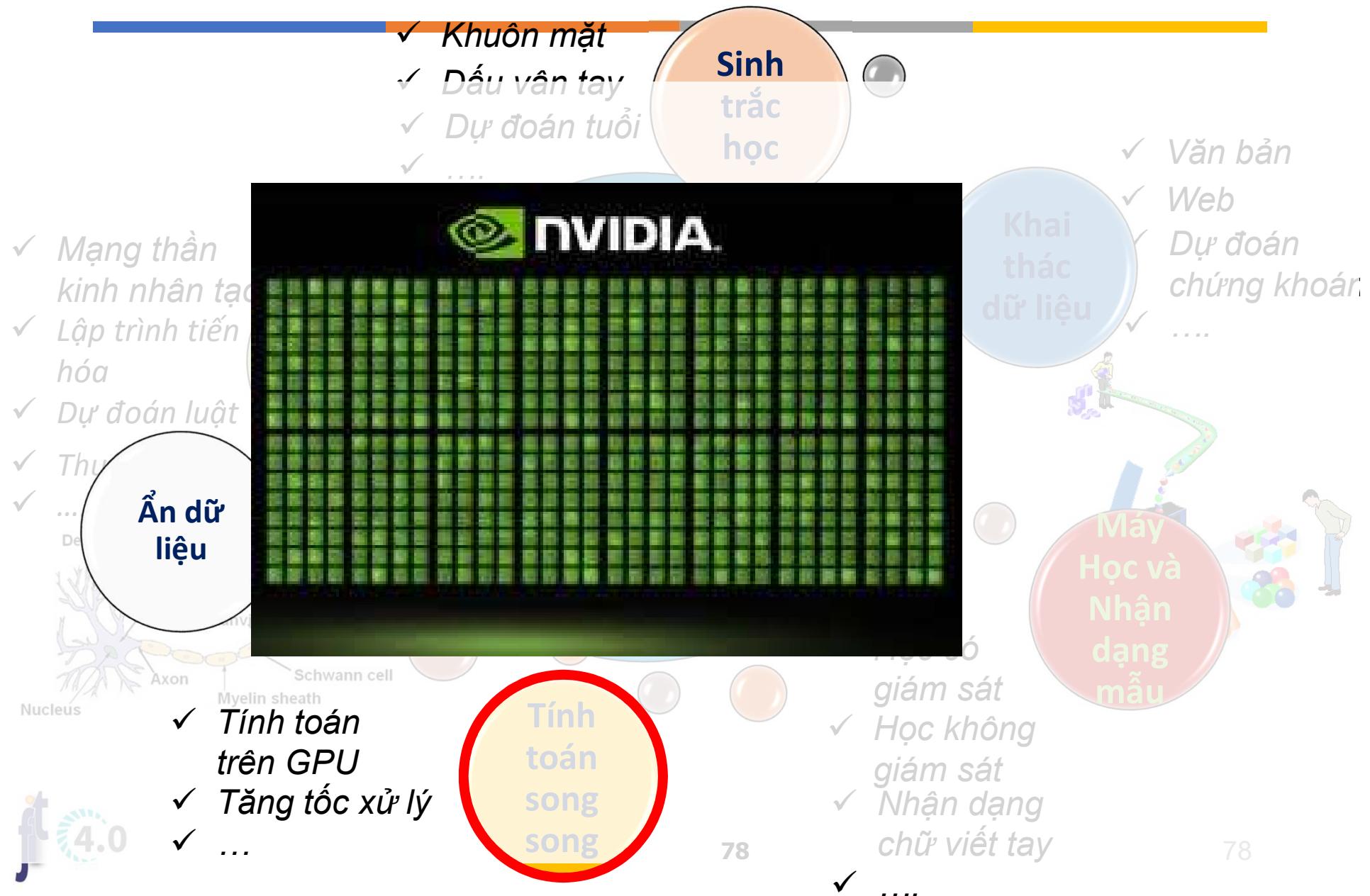
Tính toán song song

- ✓ Văn bản
- ✓ Web
- ✓ Dự đoán
- ✓ Chứng khoán
-

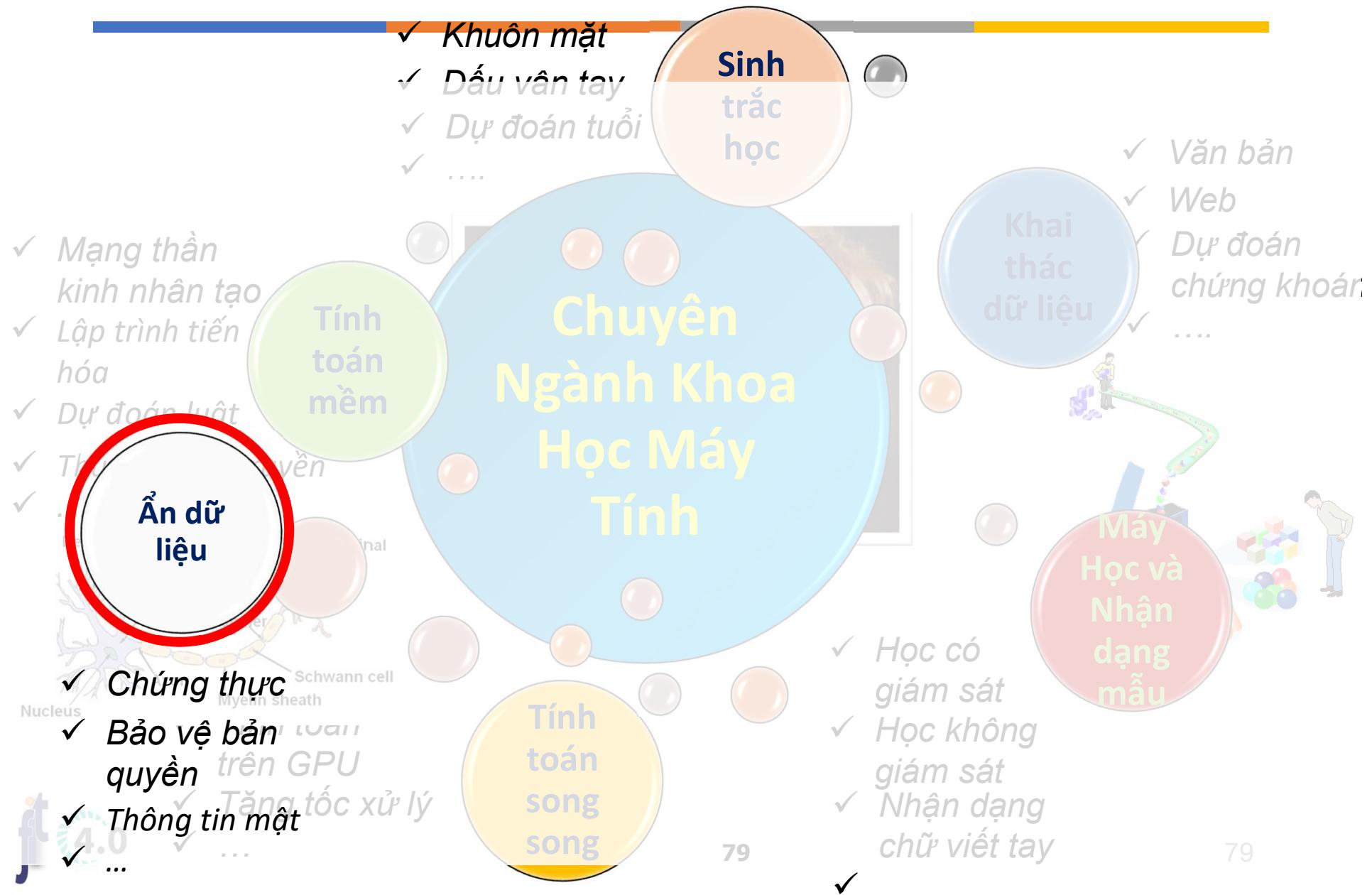
Mày
Học và
Nhận
dạng
mẫu

- ✓ Học có giám sát
- ✓ Học không giám sát
- ✓ Nhận dạng chữ viết tay
-

Hướng nghiên cứu ngành



Hướng nghiên cứu ngành



Các học phần chung chuyên ngành

- SV tích lũy ít nhất 5 học phần trong danh sách

Cơ sở trí tuệ nhân tạo

Automat và ngôn ngữ hình thức

Các hệ cơ sở tri thức

Đồ họa máy tính

Khai thác dữ liệu và ứng dụng

Máy học

Nhận dạng

Nhập môn phân tích độ phức
tạp thuật toán

Xử lý ảnh số và video số

Nhập môn phân tích và thiết kế
giải thuật



Các học phần tự chọn ngành

- SV tích lũy ít nhất 5 học phần tự chọn, trong đó ít nhất có 2 HP tương đương 8 tín chỉ thuộc ngành KHMT sau:

Trình biên dịch

Kĩ thuật trí tuệ nhân tạo

Ẩn dữ liệu và chia sẻ thông tin

Lập trình nhúng cơ bản

Kiến tập nghề nghiệp (3 tín chỉ)

Lập trình nhúng nâng cao

Khởi nghiệp (2 tín chỉ)

Lập trình song song trên GPU

Các học phần tự chọn ngành (tt)

Logic mờ và ứng dụng

Nguyên lý các ngôn ngữ lập trình

Sinh trắc học

Cơ sở tính toán và truyền thông

Phương pháp luận sáng tạo

Khai thác dữ liệu nâng cao

Khoa học về Web

Hệ thống thông minh

Phương pháp nghiên cứu khoa học

Và các học phần cơ sở và chuyên sâu của các chuyên ngành khác

Định hướng nghề nghiệp

Tác nhân thông minh

- Kỹ sư phát triển Game, Kỹ sư điều khiển tự động

Phân tích dữ liệu

- Kỹ sư phân tích và dự đoán xu hướng

Bản quyền dữ liệu

- Kỹ sư phần mềm hỗ trợ chứng thực tác quyền

Tính toán song song

- Kỹ sư phát triển Game và các ứng dụng tính toán hiệu năng cao

Nghiên cứu cải tiến và khám phá ý tưởng mới

- Nhà nghiên cứu

Một số hướng nghiên cứu đang thực hiện (2014-2016)

- Khai thác ý kiến xã hội
 - Nghiên cứu thuật toán và giải pháp công nghệ để đánh giá ý kiến người dùng trên mạng xã hội.
 - Trưởng nhóm: PGS.TS.Lê Hoài Bắc
 - Thành viên:
 - TS.Tô Hoài Việt
 - ThS.Lê Ngọc Thành
 - ThS.Bùi Thị Danh
 - ThS.Nguyễn Tiến Huy

Một số hướng nghiên cứu đang thực hiện (2014-2016)

- Các bài toán liên quan đến di chuyển trên xa lộ
 - Nghiên cứu thuật toán và giải pháp công nghệ để ứng dụng vào giao thông công cộng như phát hiện tài xế ngủ gật, tư vấn lộ trình, tư vấn giá vận chuyển,...
 - Trưởng nhóm: TS.Lê Mai Tùng (đang công tác tại Úc)
 - Thành viên:
 - ThS.Lê Ngọc Thành

THÔNG TIN, DỮ LIỆU, VÀ TRI THỨC

Tổng quan

- Máy tính có thể biểu diễn bất kỳ thông tin nào nếu đã được rời rạc hóa và số hóa
- Thuật toán có thể xử lý/chuyển đổi các *thông tin số* bằng nhiều cách khác nhau
 - Xáo trộn *dữ liệu số* từ phần này sang phần khác của máy tính
 - Tìm kiếm các dạng mẫu, tạo mô phỏng, làm tương quan các kết nối để tạo ra *tri thức* mới và hiểu biết mới

Thuật ngữ liên quan

- Thông tin (information)
 - Là tất cả những gì đem lại sự hiểu biết về thế giới xung quanh
- Dữ liệu (data)
 - Là hình thức thể hiện của thông tin với mục đích lưu trữ/xử lý nào đó
- Tri thức (knowledge)
 - Là mục đích của nhận thức trên cơ sở tiếp nhận thông tin
 - Quá trình xử lý thông tin = quá trình nhận thức để có tri thức

Câu hỏi đặt ra

- Máy tính lưu trữ dữ liệu số như thế nào?
 - Số, văn bản, hình ảnh, âm thanh, và video
- Làm thế nào mà máy tính có thể xắp xỉ dữ liệu tương tự trong thế giới thực?
- Làm sao máy tính có thể phát hiện và ngăn chặn lỗi trong dữ liệu?

CƠ SỞ DỮ LIỆU

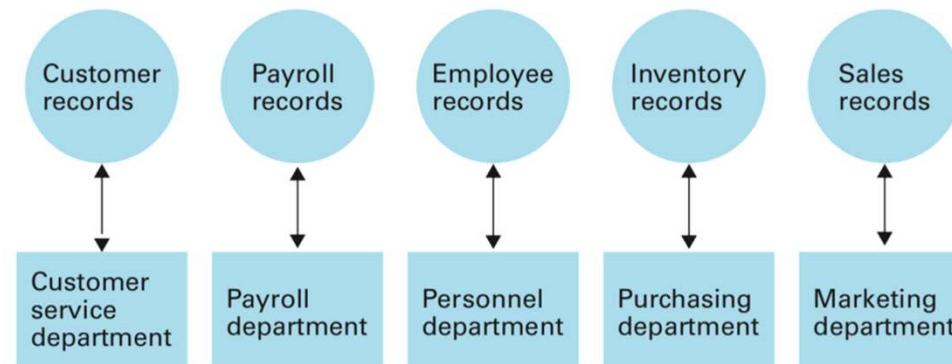
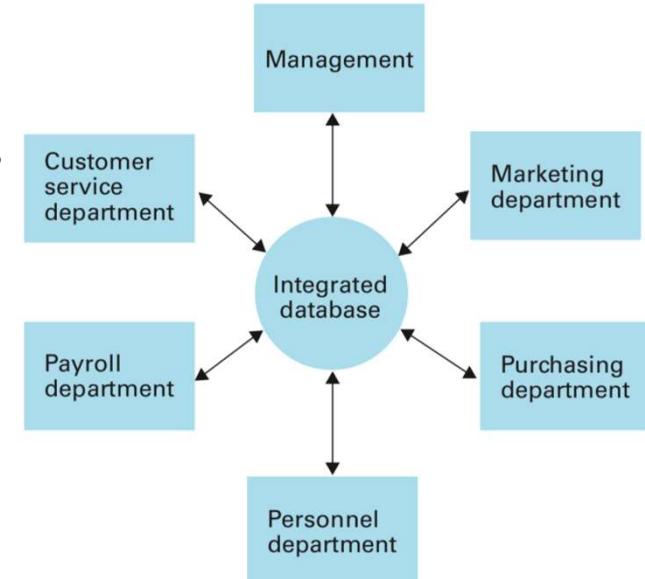
Định nghĩa CSDL

- Cơ sở dữ liệu

- cơ sở dữ liệu là tập dữ liệu đa chiều, các liên kết nhiều khía cạnh với nhau.

- Hệ thống file

- Dữ liệu một chiều thể hiện thông tin từ một khía cạnh duy nhất

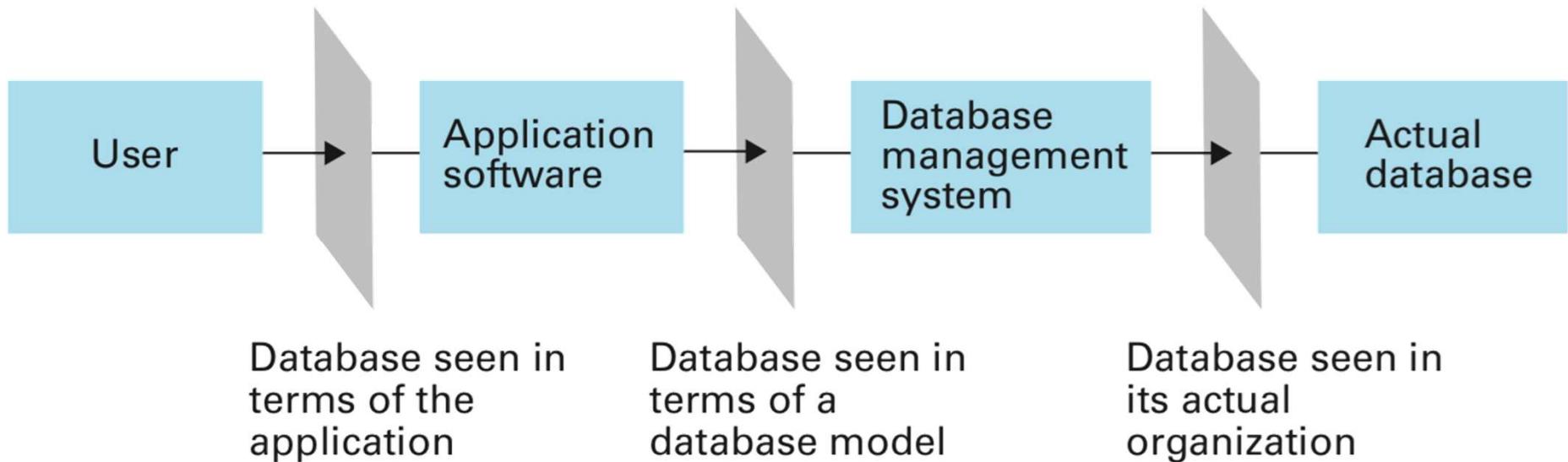


From Brooks, 2015

Hệ quản trị CSDL

- Hệ quản trị CSDL (DBMS)
 - Một lớp phần mềm hỗ trợ các thao tác với cơ sở dữ liệu để đáp ứng các yêu cầu từ các ứng dụng
- Cơ sở dữ liệu phân tán
 - Một cơ sở dữ liệu được lưu trữ trên nhiều máy
- Độc lập dữ liệu
 - Khả năng thay đổi cấu trúc cơ sở dữ liệu mà không thay đổi phần mềm ứng dụng sử dụng nó

Triển khai ứng dụng cơ sở dữ liệu



From Brockshear, 2015

Mô hình cơ sở dữ liệu

- Góc nhìn cơ sở dữ liệu:
 - Mô hình quan hệ
 - Mô hình hướng đối tượng
 - Mô hình phân cấp

Mô hình CSDL quan hệ

- Cơ sở dữ liệu là tập hợp các bảng
- Mỗi bảng được gọi là một mối quan hệ
- Mỗi cột trong bảng ghi lại một thuộc tính
- Một hàng trong bảng được gọi là một tuple

NGÀNH HỆ THỐNG THÔNG TIN

Tổng quan

BỘ MÔN HỆ THỐNG THÔNG TIN

10



- Thành lập năm 1995, cùng với việc thành lập khoa CNTT
- Địa chỉ: Phòng 184, 227 Nguyễn Văn Cừ, Q5 TpHCM
- Điện thoại: (08) 38 309 291 (ext: 801)

Nguồn nhân lực

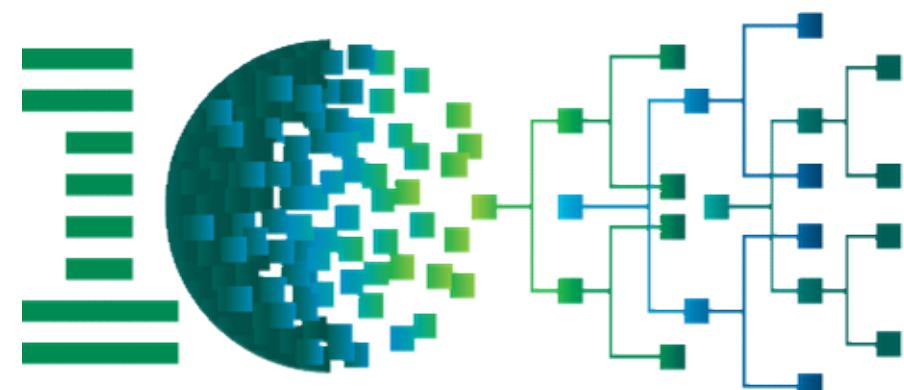
- 18 giảng viên:
 - 8 Tiến sĩ (có 1 Phó Giáo sư)
 - 4 NCS (1: Czech, 1: Pháp, 1: Mỹ, 1: Ý)
 - 6 Thạc sĩ

Trưởng bộ môn: TS. Phạm Nguyễn Cương
• **Phó Trưởng bộ môn: TS. Nguyễn Trần Minh Thư**



HTTT là gì

Là tập các thành phần cùng nhau thu thập, xử lý, lưu trữ, phân phối thông tin nhằm hỗ trợ cho quá trình vận hành và ra quyết định của một tổ chức nghiệp vụ



From ibmims

Khái niệm *Quy trình nghiệp vụ*

- Quy trình xử lý dữ liệu/thông tin để thực hiện một nghiệp vụ cụ thể
- VD: đối với HTTT quản lý SV của một trường đại học, có quy trình nghiệp vụ đăng ký học phần của SV, quy trình nghiệp vụ theo dõi đóng học phí,...

Khái niệm Quy trình phát triển HTTT

- Gần giống quy trình phát triển phần mềm, vì một HTTT tin học hóa được thể hiện qua một phần mềm ứng dụng
- Đặc biệt chú trọng về nghiệp vụ của HTTT (không liên quan gì đến tin học) và yếu tố người dùng cuối (là những người thực hiện nghiệp vụ, như nghiệp vụ quản lý đăng ký học phần của SV)

Khái niệm Công cụ PT HTTT

- Công cụ hỗ trợ giai đoạn thu thập và phân tích nhu cầu
- Công cụ hỗ trợ giai đoạn phân tích dữ liệu và quy trình nghiệp vụ
- Công cụ hỗ trợ giai đoạn thiết kế và hiện thực (lập trình) HTTT
- Công cụ hỗ trợ kiểm thử

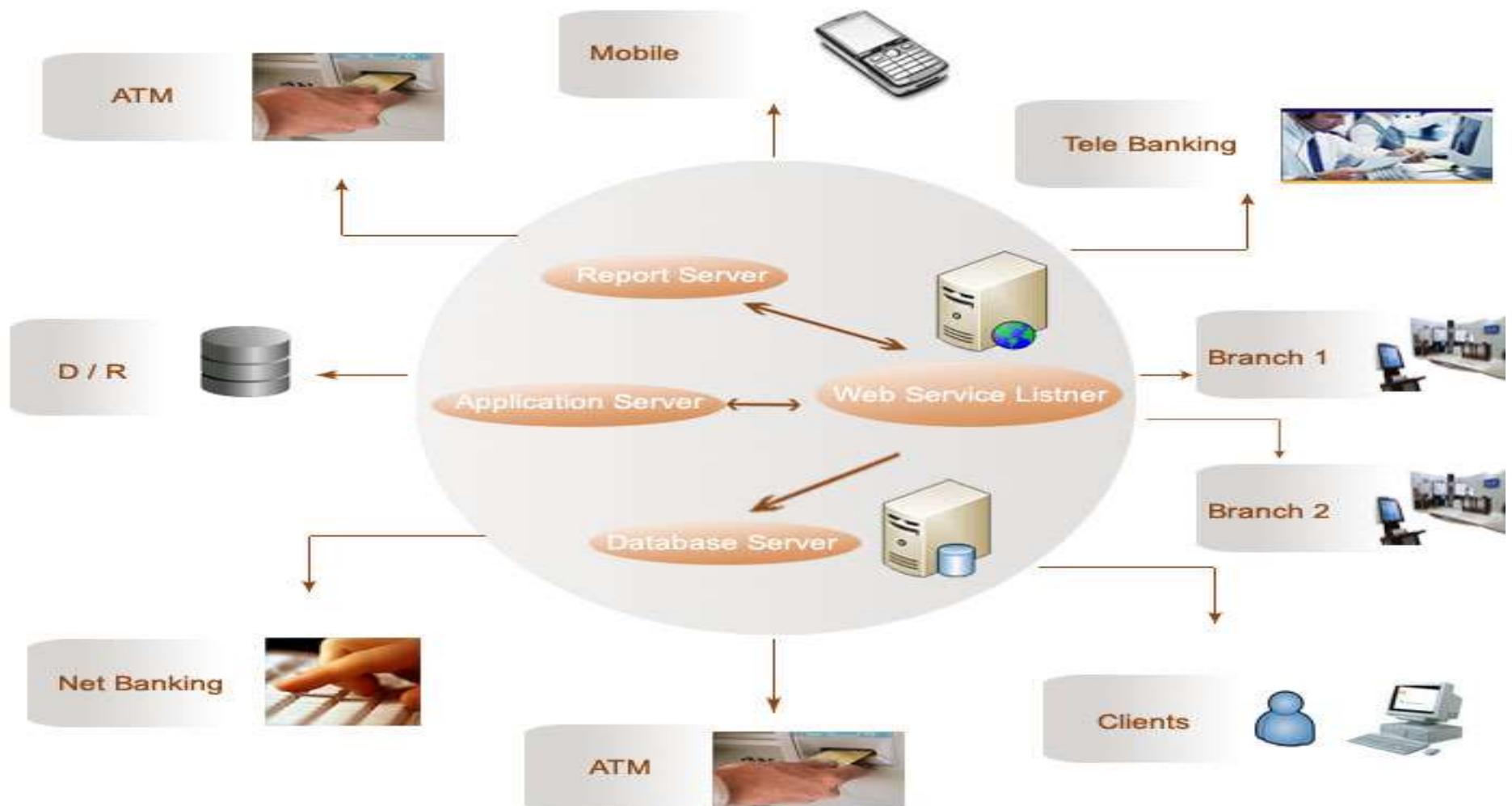
Mục tiêu đào tạo

- Khả năng *phân tích yêu cầu, thiết kế, cài đặt, kiểm thử và triển khai thực hiện các dự án hệ thống thông tin*
- Khả năng *sử dụng và tự tìm hiểu để sử dụng công cụ hỗ trợ phát triển hệ thống thông tin (công cụ hỗ trợ phân tích/thiết kế, công cụ phát triển phần mềm: môi trường mạng/ phân tán/ di động)*
- Khả năng *tự tìm hiểu, nghiên cứu các công nghệ, phương pháp, quy trình mới trong lĩnh vực hệ thống thông tin*

Hệ thống
thông tin là gì?



Ví dụ mô hình HTTT ngân hàng



Chức năng HTTT

ACB NGÂN HÀNG Á CHÂU
Ngân hàng của mọi nhà

HỖ TRỢ KHÁCH HÀNG [24/7]
08 38 247 247
1800 577 776 (miễn phí)

Thống tin tài khoản
Thông tin thẻ
Liệt kê giao dịch
Chuyển khoản, chuyển tiền
Đăng ký TK thu hướng
Thanh toán hóa đơn
Bán ngoại tệ cho ACB
Mở TK đầu tư trực tuyến
Tài khoản tiền gửi có kỳ hạn
Thông Báo Tự Động
Đăng ký vay qua mạng
Đăng ký làm thẻ tín dụng
Tra cứu thông tin
Thông tin chứng khoán
Mobile Banking
Tiện ích
Trang chủ

TSAN NGOC NGAN
Thoát

fit 4.0

Để xây dựng
Hệ thống thông tin
cần làm gì?



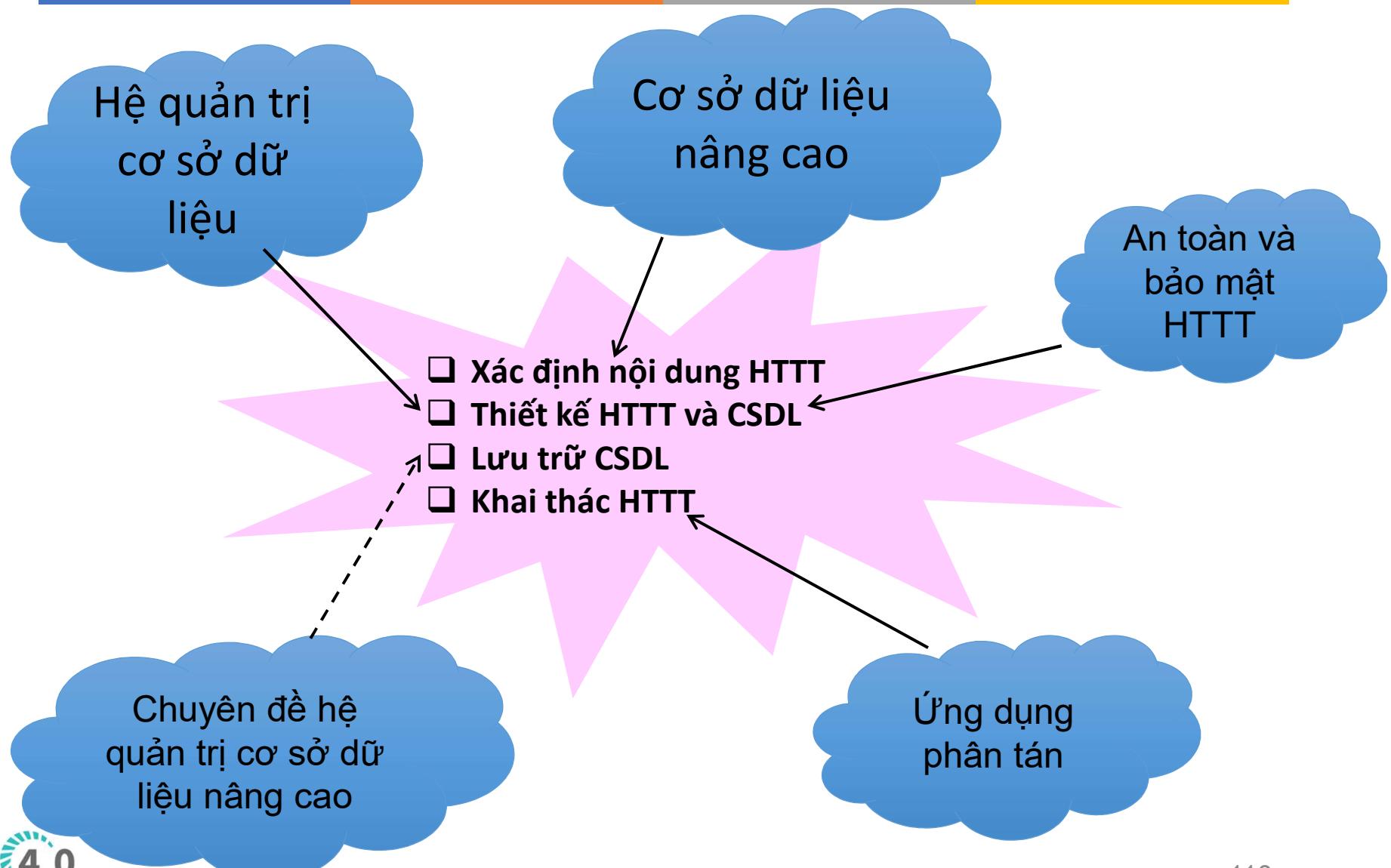
- Xác định nội dung HTTT
- Thiết kế HTTT và CSDL
- Lưu trữ CSDL
- Khai thác HTTT

Thiết kế HTTT ?

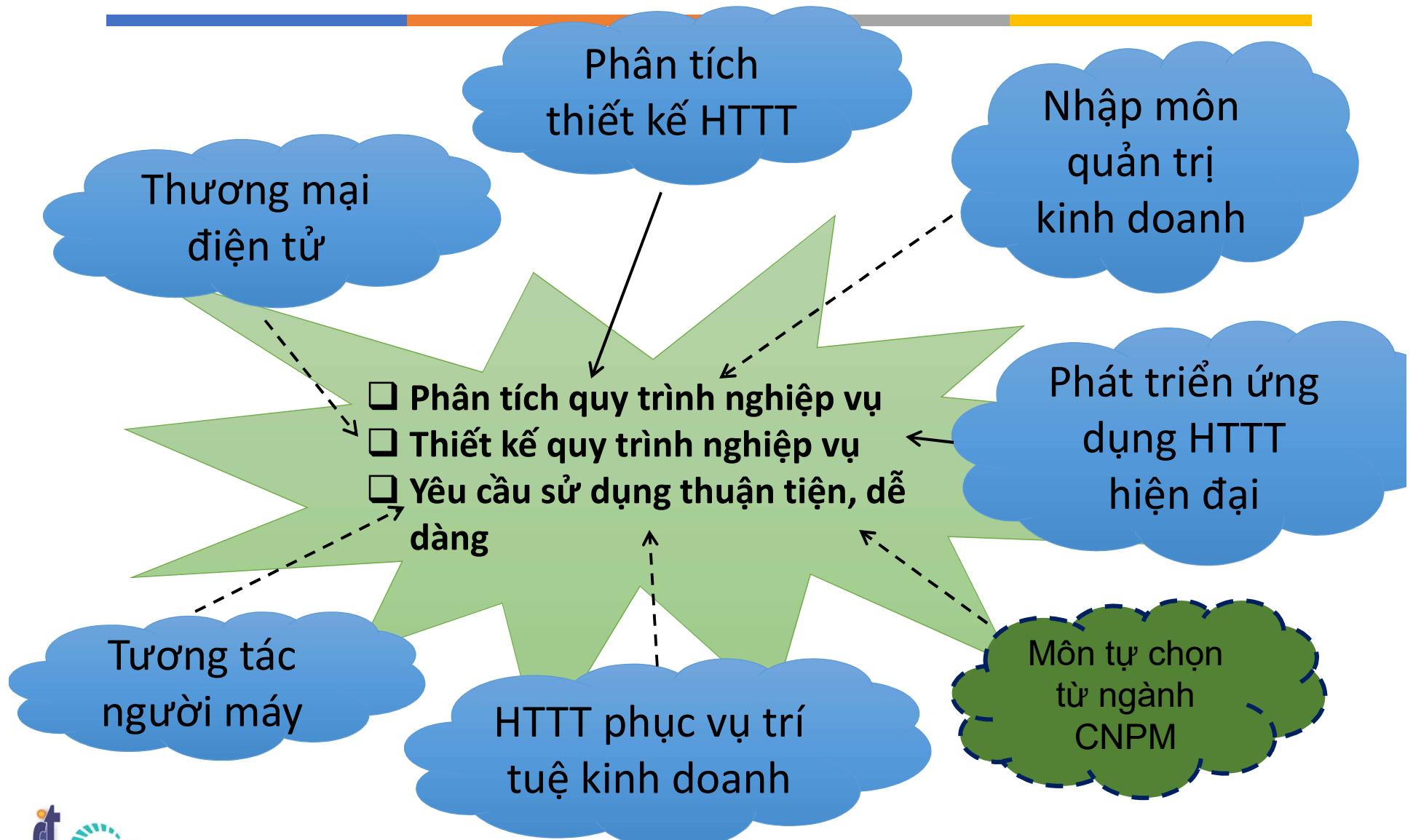


- Phân tích quy trình nghiệp vụ ngoài thực tế
- Thiết kế quy trình nghiệp vụ trong môi trường công nghệ
- Yêu cầu sử dụng thuận tiện, dễ dàng
-

Các học phần liên quan:



Các học phần liên quan:



Các học phần bắt buộc của nhóm ngành CNTT & ngành HTTT

Cơ sở dữ liệu

- SV tích lũy ít nhất 5 học phần bắt buộc của
ngành HTTT:

Cơ Sở dữ liệu nâng cao

Hệ quản trị CSDL

Phân tích thiết kế HTTT

Phát triển ứng dụng HTTT hiện
đại

An toàn và bảo mật trong HTTT

Các học phần tự chọn ngành

- SV tích lũy ít nhất 5 học phần tự chọn, trong đó ít nhất có 2 HP tương đương 8 tín chỉ thuộc ngành HTTT sau:

Tương tác người máy

Chuyên đề hệ QT cơ sở
dữ liệu nâng cao

Thương mại điện tử

Nhập môn Quản trị
doanh nghiệp

Ứng dụng phân tán

HTTT phục vụ trí tuệ
kinh doanh

Chuyên đề chọn lọc
trong HTTT

Biểu diễn thông tin

Kiến tập nghề nghiệp (3 tín chỉ)

Khởi nghiệp (2 tín chỉ)

Một số HP tự chọn thuộc ngành khác

- **Ngành KTPM:**

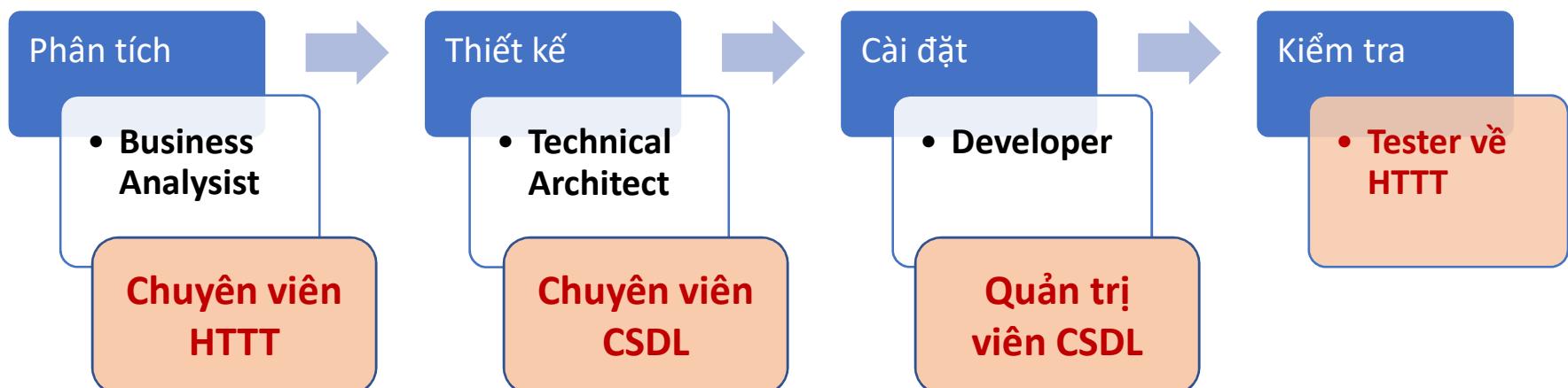
- Phân tích và thiết kế phần mềm
- Phân tích và quản lý yêu cầu phần mềm
- Phát triển phần mềm cho thiết bị di động
- Lập trình ứng dụng Java
- Công nghệ XML và ứng dụng
- Các công nghệ mới trong phát triển phần mềm
- Lập trình hướng đối tượng nâng cao

- **Ngành KHMT:**

- Khai thác dữ liệu và ứng dụng

Định hướng nghề nghiệp

Trong công nghệ phần mềm và HTTT:



- Team Leader
- Project Manager về dự án HTTT
- Consultant về HTTT & CSDL

Hướng nghiên cứu ngành

- Các hướng gắn liền với CSDL
- Các hướng nghiên cứu gắn liền với Internet, thiết bị di động, mạng điện rộng: truy tìm thông tin, biểu diễn thông tin, xử lý tự động thông tin đa truyền thông, an toàn và bảo mật,...

Hướng nghiên cứu ngành (tt)

• Hệ thống xử lý văn bản:

- Các chủ đề cơ bản: phân lớp văn bản, gom nhóm văn bản, tìm kiếm văn bản, rút trích thông tin trong văn bản.
- Các chủ đề nâng cao: hệ thống hỏi đáp văn bản (question and answer system), hệ thống suy luận văn bản (textual entailment system), hệ thống rút trích thông tin văn bản y tế (medical information extraction)

Chủ trì: PGS. TS .Hồ Bảo Quốc

- Tham gia: TS. Nguyễn Trường Sơn; TS. Lê Nguyễn Hoài Nam; NCS. Trần Quang Vinh; ThS. Tuấn Nguyễn Hoài Đức

Hướng nghiên cứu ngành (tt)

- HTTT di động:

- Mô hình và giải pháp phục vụ truy vấn hiệu quả trong hệ thống dữ liệu di động

Chủ trì: TS. Nguyễn Trần Minh Thư

- ❑ Tham gia: NCS. Hồ Lê Thị Kim Nhung;
ThS. Tiết Gia Hồng; Ths. Phạm Minh Tú

Hướng nghiên cứu ngành (tt)

- Hệ thống tư vấn

- Hệ thống tư vấn tour du lịch
- Hệ thống tư vấn phim
- Hệ thống tư vấn đăng ký học phần
- Đánh giá hệ thống tư vấn

Chủ trì: Ths. Hồ Thị Hoàng Vy.

Hướng nghiên cứu ngành (tt)

- **Bảo mật cơ sở dữ liệu:**
 - Bảo vệ tính riêng tư khi khai thác CSDL đặt ở nhà cung cấp dịch vụ
- Chủ trì: TS. Phạm Thị Bạch Huệ
- Tham gia: ThS. Lương Vĩ Minh, NCS. Hoàng Anh Tú

Hướng nghiên cứu ngành (tt)

- Khai thác thông tin y sinh:
 - Học các motif chuyên biệt để dự đoán viêm gan C (Discriminative Motif Learning for Hepatitis C Virus Study)

Chủ trì: TS. Lê Thị Nhàn

Hướng nghiên cứu ngành (tt)

- **Dịch vụ thông minh trong môi trường giáo dục và kinh doanh:**

- Hệ thống hỗ trợ học tập dựa trên ngữ cảnh
- Chatbot

Chủ trì: TS. Phạm Nguyễn Cường

□ Tham gia: ThS. Tiết Gia Hồng, ThS. Hồ Lê Thị Kim Nhung, TS. Lê Nguyễn Hoài Nam

Hướng nghiên cứu ngành (tt)

- **Ứng dụng tin học trong giáo dục:**
 - Hệ thống hỗ trợ dành cho việc học với ứng dụng
Chủ trì: TS. Thái Lê Vinh
 - Tham gia: NCS. Vũ Thị Mỹ Hằng