

TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN

KHOA HỆ THỐNG THÔNG TIN

Tài liệu bài giảng:

MÔN CƠ SỞ DỮ LIỆU

Chương 1:

TỔNG QUAN

ThS. Dương Phi Long – Email: longdp@uit.edu.vn

NỘI DUNG BÀI HỌC

01



Các khái niệm của một hệ CSDL

02



Các mô hình dữ liệu

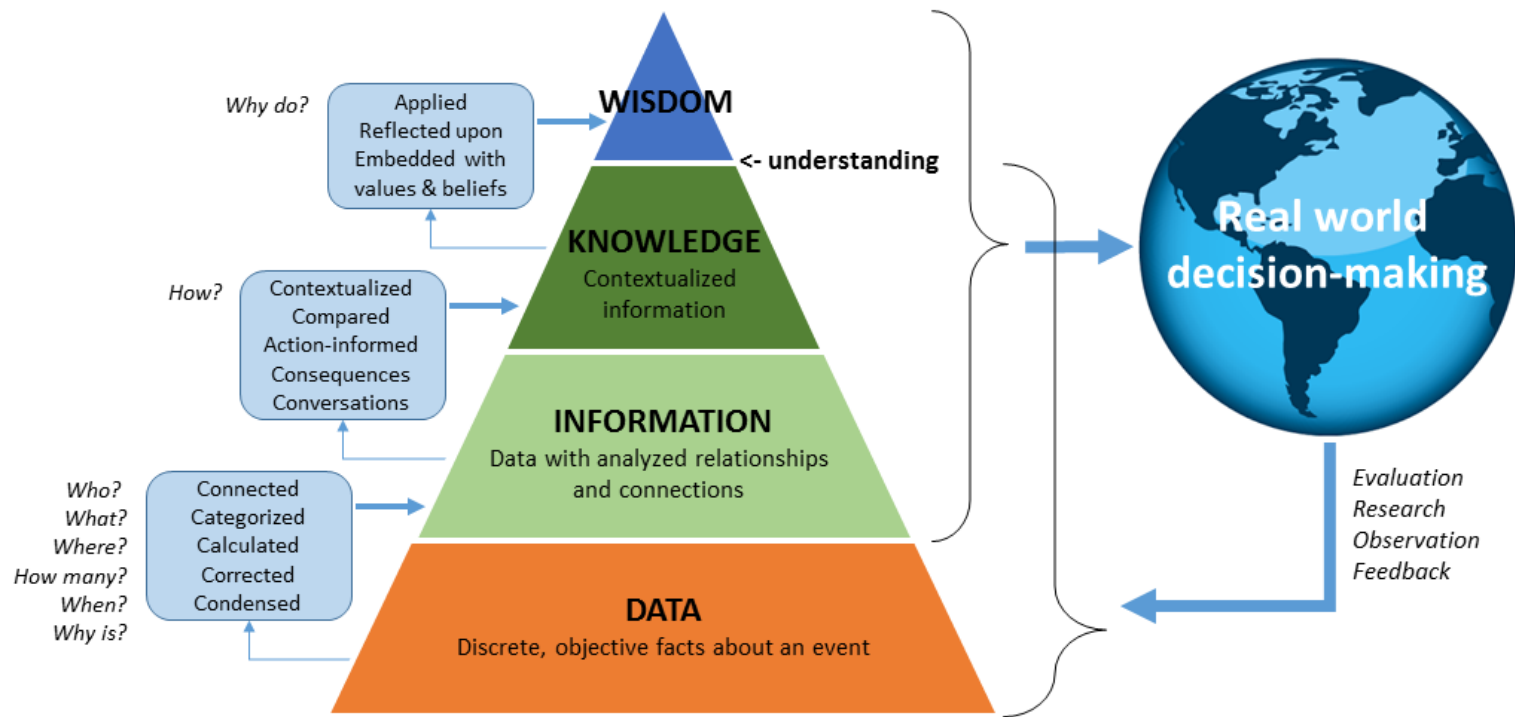
1

Các khái niệm của một hệ CSDL

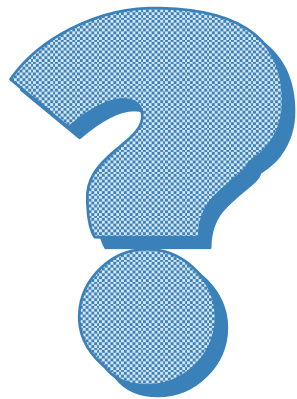
1. Đặt vấn đề
2. Dữ liệu (Data)
3. Hệ thống tập tin (File System)
4. Định nghĩa một CSDL
5. Các đối tượng sử dụng CSDL
6. Hệ quản trị CSDL
7. Các mức của một CSDL

[illegible]

1. Đặt vấn đề



1. Đặt vấn đề



Vấn đề 01

Quản lý dữ liệu như thế nào?



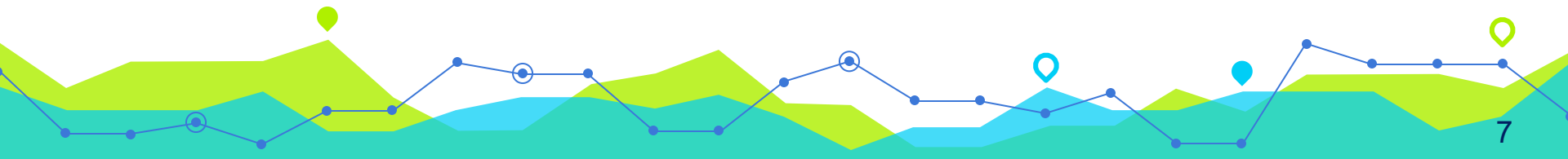
Vấn đề 02

Cơ sở dữ liệu là gì?



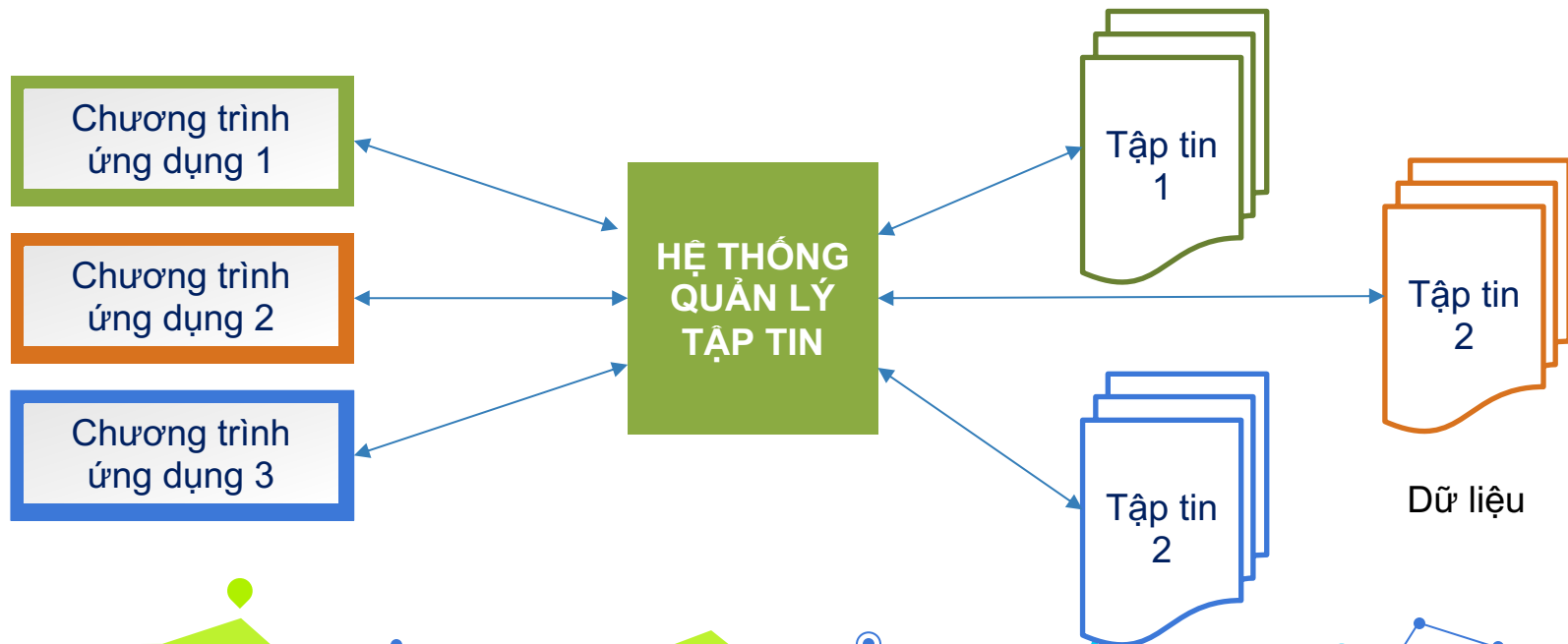
2. Dữ liệu (Data)

- Một mô tả hình thức về những sự kiện, khái niệm
- Ví dụ:
 - Thông tin cá nhân của SV
 - Bảng điểm của SV
 - Hóa đơn bán hàng
 - Báo cáo doanh thu bán hàng



2. Hệ thống tập tin (File system)

Là tập hợp các tập tin riêng lẻ phục vụ cho mục đích của đơn vị sử dụng.



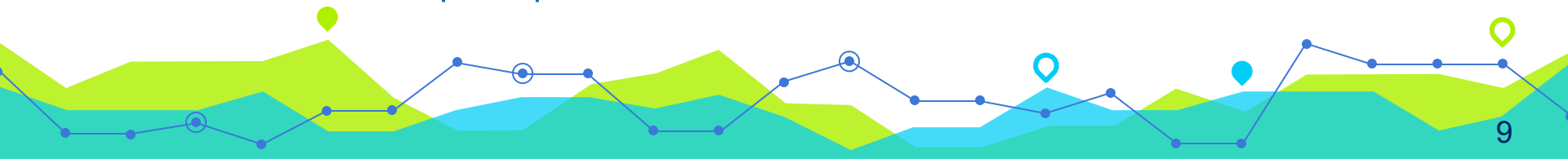
2. Hệ thống tập tin (File system)

- Ưu điểm:

- Triển khai ứng dụng nhanh
- Khả năng đáp ứng nhanh chóng, kịp thời (vì chỉ phục vụ cho mục đích hạn hẹp)

- Nhược điểm:

- Trùng lặp dữ liệu → lãng phí, dữ liệu không nhất quán
- Khó khăn trong việc truy xuất thông tin
- Chi phí cao
- Chia sẻ dữ liệu hạn chế



3. Cơ sở dữ liệu (Database)

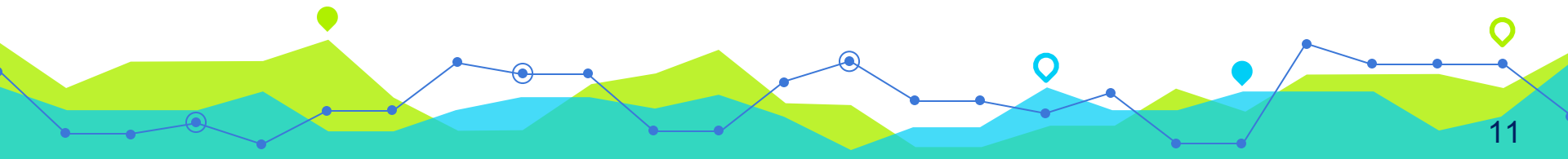
- Là một hệ thống các thông tin có cấu trúc, được lưu trữ trên các thiết bị lưu trữ nhằm thỏa mãn yêu cầu khai thác thông tin đồng thời của nhiều người sử dụng hay nhiều chương trình ứng dụng với những mục đích khác nhau



3. Cơ sở dữ liệu (Database)

- Ưu điểm:

- Giảm trùng lặp thông tin xuống mức thấp nhất, đảm bảo tính nhất quán và toàn vẹn dữ liệu.
- Đảm bảo dữ liệu được truy xuất theo nhiều cách khác nhau.
- Khả năng chia sẻ thông tin cho nhiều người, nhiều ứng dụng khác nhau.



3. Cơ sở dữ liệu (Database)

- Một số vấn đề cần giải quyết:

- Dữ liệu phải được cập nhật thường xuyên
- Tính bảo mật và quyền khai thác thông tin của người sử dụng
- Tranh chấp dữ liệu (xử lý đồng thời)
- Đảm bảo dữ liệu khi có sự cố (sao lưu, phục hồi)



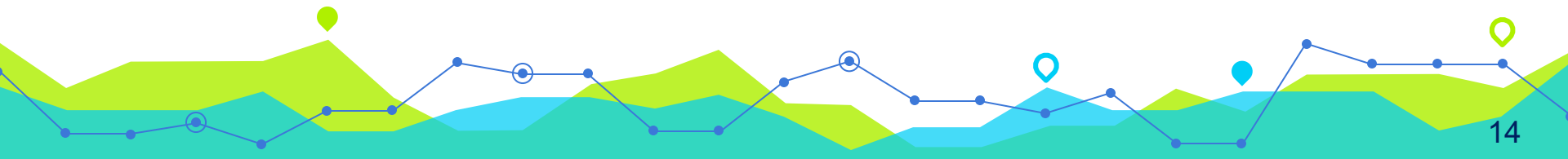
4. Các đối tượng sử dụng CSDL

- **Người dùng cuối (End user):**
 - Không chuyên về lĩnh vực tin học và CSDL
 - Cần công cụ để khai thác CSDL (y tế, trường học, kinh doanh, ...)
- **Chuyên viên tin học:** Xây dựng các ứng dụng để phục vụ cho các mục đích quản lý.

4. Các đối tượng sử dụng CSDL

- Quản trị viên (Database Administrator):

- Tổ chức CSDL
- Xử lý bảo mật, cấp quyền CSDL
- Sao lưu, phục hồi dữ liệu
- Giải quyết các tranh chấp dữ liệu



5. Hệ quản trị CSDL

- **Hệ quản trị CSDL (DBMS – Database Management System) :**
 - Là tập hợp các chương trình cho phép người sử dụng **tạo ra và duy trì CSDL**.
 - Là phần mềm cho phép **định nghĩa, xây dựng và xử lý dữ liệu**:
 - Định nghĩa – khai báo bộ khung dữ liệu, mô tả chi tiết về dữ liệu
 - Xây dựng – lưu trữ dữ liệu
 - Xử lý – Truy vấn, cập nhật dữ liệu và truy xuất báo cáo
 - Là hệ thống các phần mềm hỗ trợ tích cực cho các nhà phân tích, thiết kế và khai thác CSDL.

5. Hệ quản trị CSDL

- Một số DBMS thông dụng:

- Microsoft SQL Server
- My SQL
- Oracle
- Visual FoxPro
- Microsoft Access
- DB2

→ Nhiều DBMS hiện nay dựa trên
mô hình dữ liệu quan hệ.

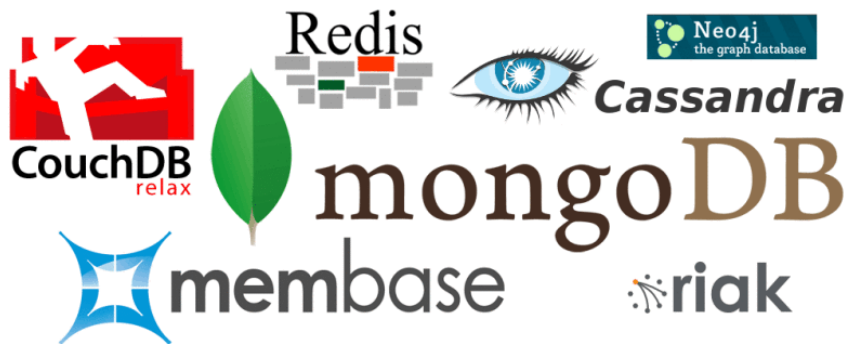


ORACLE®

5. Hệ quản trị CSDL

- Một số DBMS khác:

- NoSQL (Redis, Cassandra, MongoDB, Neo4j, ...)
- NewSQL (Apache Trafodion, CockroachDB, VoltDB, NuoDB...)

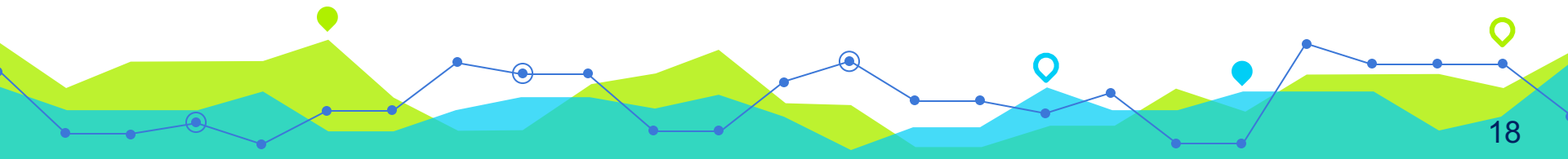


5. Hệ quản trị CSDL

- Một DBMS phải có:

- Ngôn ngữ giao tiếp giữa người sử dụng và CSDL
- Từ điển dữ liệu (Data Dictionary)
- Có biện pháp bảo mật khi có yêu cầu
- Cơ chế giải quyết tranh chấp dữ liệu
- Có cơ chế sao lưu (backup), phục hồi (restore)
- Đảm bảo tính độc lập giữa dữ liệu và chương trình

→ Mỗi DBMS khác nhau có cơ chế khác nhau để giải quyết các vấn đề



5. Hệ quản trị CSDL

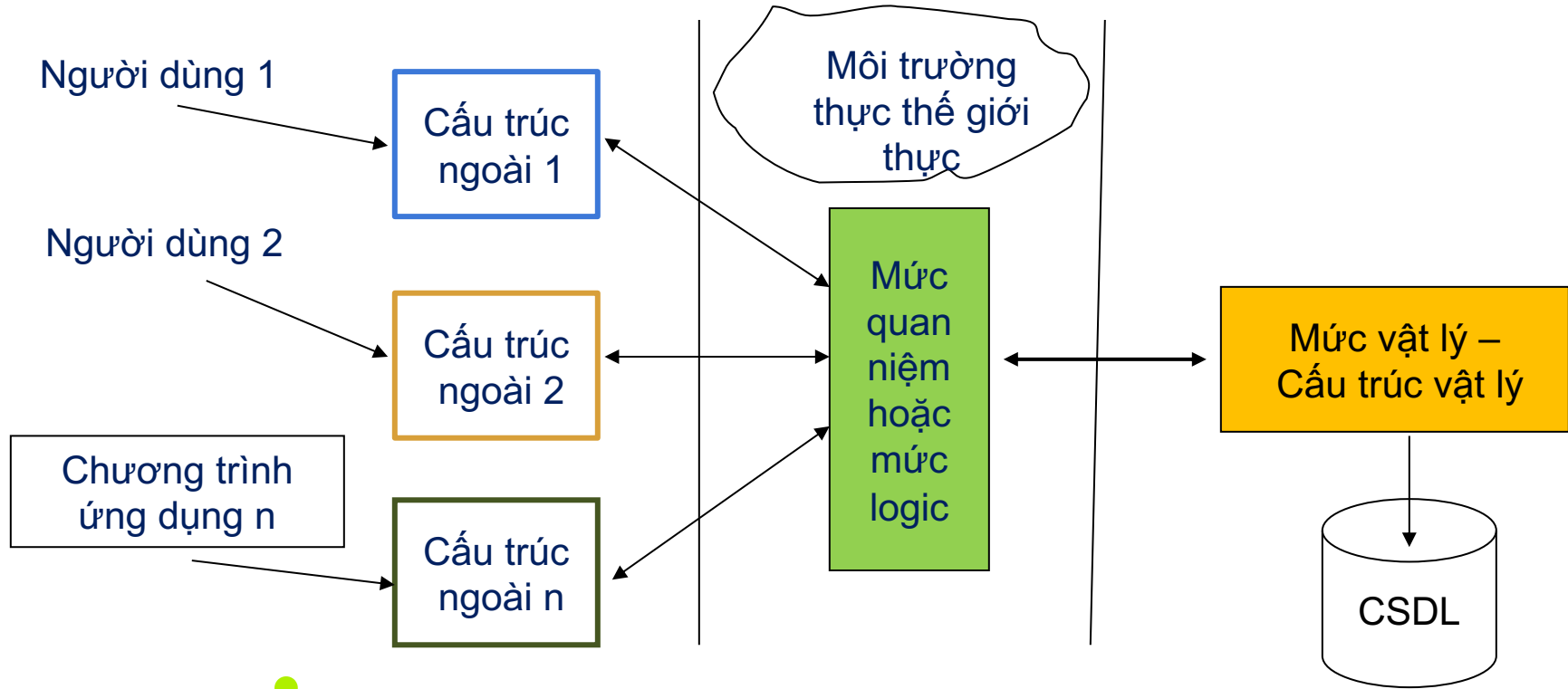
- **Ngôn ngữ giao tiếp:** được chia thành 04 loại
 - **Ngôn ngữ mô tả dữ liệu** (DDL–Data Definition Language): cho phép khai báo
 - Cấu trúc CSDL
 - Các mối liên hệ của dữ liệu
 - Các quy định, ràng buộc dữ liệu.
 - **Ngôn ngữ thao tác dữ liệu** (DML–Data Manipulation Language) cho phép thực hiện các thao tác Thêm, Xóa, Sửa dữ liệu

5. Hệ quản trị CSDL

- **Ngôn ngữ giao tiếp:** được chia thành 04 loại
 - **Ngôn ngữ truy vấn dữ liệu** (SQL–Structured Query Language): cho phép người khai thác sử dụng để truy vấn thông tin cần thiết.
 - **Ngôn ngữ quản lý dữ liệu** (DCL–Data Control Language): cho phép quản lý bảo mật, phân quyền cho người sử dụng.



6. Các mức biểu diễn một CSDL



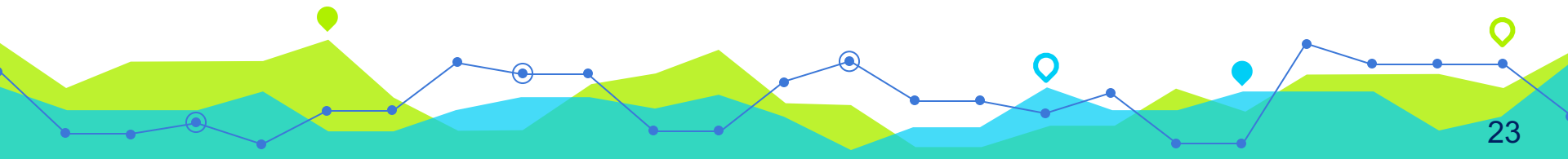
6. Các mức biểu diễn một CSDL

- **Mức trong:** (mức vật lý – Physical) là mức lưu trữ CSDL
 - Cần giải quyết vấn đề gì?
 - Dữ liệu gì?
 - Lưu trữ như thế nào? ở đâu?
 - Cần các chỉ mục gì?
 - Truy xuất tuần tự hay ngẫu nhiên.
 - Dành cho người quản trị CSDL.



6. Các mức biểu diễn một CSDL

- **Mức quan niệm:** (mức Conception hay mức Logical)
 - Cần phải lưu trữ bao nhiêu loại dữ liệu?
 - Là dữ liệu gì? mối quan hệ giữa chúng?
- **Mức ngoài:** của người sử dụng và các chương trình ứng dụng



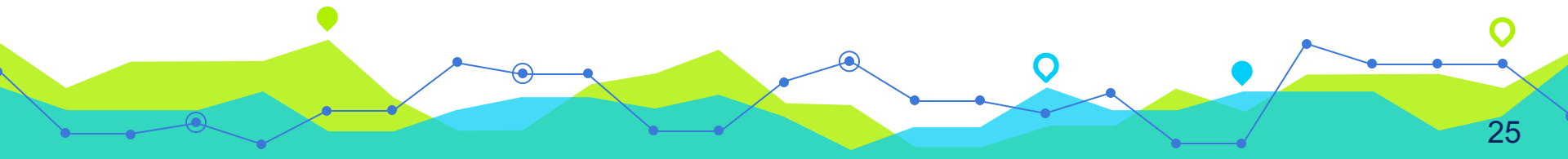


2 Các mô hình dữ liệu

1. Giới thiệu
2. Mô hình dữ liệu mạng
3. Mô hình dữ liệu phân cấp
4. Mô hình thực thể mối kết hợp
5. Mô hình dữ liệu quan hệ
6. Mô hình dữ liệu hướng đối tượng

1. Giới thiệu

- Mô hình dữ liệu là sự trừu tượng hóa của môi trường thực, biểu diễn dữ liệu ở mức quan niệm.
- **Một số mô hình:**
 - Mô hình dữ liệu mạng
 - Mô hình dữ liệu phân cấp
 - Mô hình dữ liệu thực thể mối kết hợp
 - Mô hình dữ liệu quan hệ
 - Mô hình dữ liệu hướng đối tượng



2. Mô hình dữ liệu mạng

- Được xây dựng bởi Honeywell vào năm 1964-1965
- Mô hình dữ liệu mạng (Network Data Model) còn gọi tắt là mô hình mạng (Network Model) là mô hình được **biểu diễn bởi một đồ thị có hướng**, với các thành phần:
 - Mẫu tin (record)
 - Loại mẫu tin
 - Loại liên hệ (set type)
 - Bản số

2. Mô hình dữ liệu mạng

- **Mẫu tin (record):**

- Mô tả 1 đối tượng trong thế giới thực
- **VD:** ('NV001', 'Nguyen Lam', 'Nam', '10/10/1970', 'Dong Nai')

- **Loại mẫu tin:**

- Là 1 tập các mẫu tin có cùng tính chất
- **VD:** nhân viên (NHANVIEN)
- Ký hiệu:

NHANVIEN

CONGVIEC

2. Mô hình dữ liệu mạng

- Loại liên hệ:

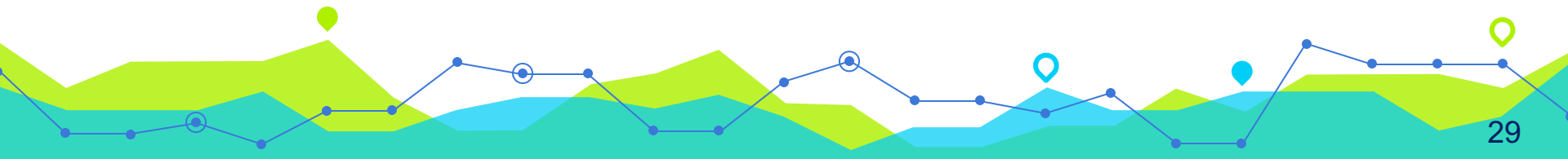
- Mô tả sự liên kết giữa 1 loại mẫu tin chủ và 1 loại mẫu tin thành viên
- VD: tham gia (ThamGia)
- Ký hiệu:

Tham gia

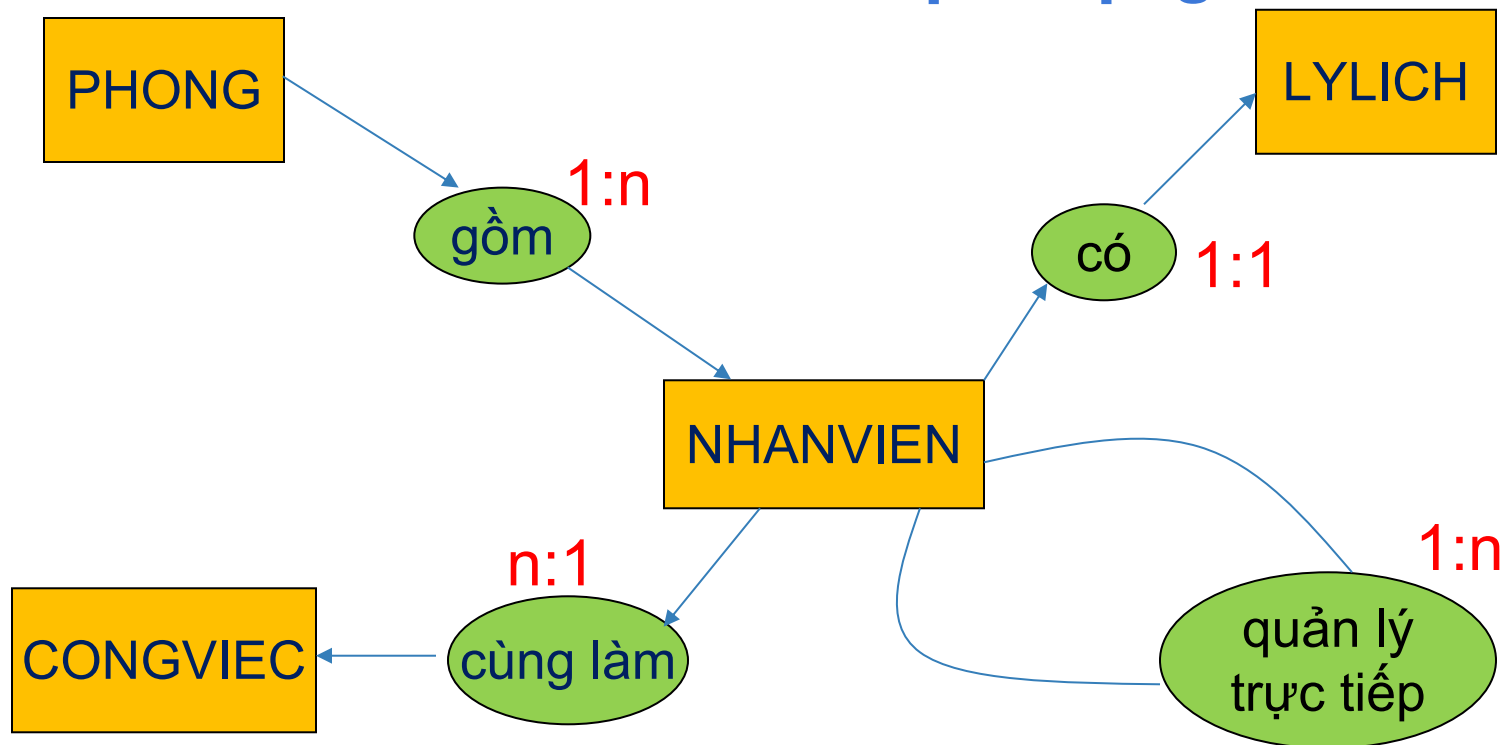
- Loại liên hệ đệ quy (Recursive): một loại mẫu tin chủ cũng có thể đồng thời là loại mẫu tin thành viên với chính nó.

2. Mô hình dữ liệu mạng

- **Bản số:** Chỉ ra số lượng các mẫu tin tham gia trong mối liên hệ
 - **(1:1)** (one-to-one): mỗi mẫu tin của loại mẫu tin chủ kết hợp với đúng 1 mẫu tin của loại mẫu tin thành viên.
 - **(1:n)** (one-to-many): mỗi mẫu tin của loại mẫu tin chủ kết hợp với 1 hay nhiều mẫu tin thành viên.
 - **(n:1)** (many-to-one): nhiều mẫu tin của loại mẫu tin chủ kết hợp với đúng 1 mẫu tin của loại mẫu tin thành viên.



2. Mô hình dữ liệu mạng



2. Mô hình dữ liệu mạng

- Nhận xét

- Tương đối đơn giản
- Dễ sử dụng
- Không thích hợp biểu diễn CSDL có quy mô lớn
- Khả năng diễn đạt ngữ nghĩa kém

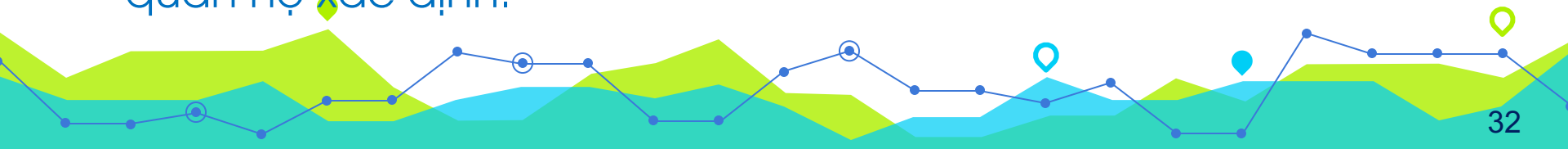


3. Mô hình dữ liệu phân cấp

- Được xây dựng bởi IBM và North American Rockwell vào khoảng năm 1965.
- Mô hình dữ liệu phân cấp (Hierarchical Data Model) còn gọi tắt là mô hình phân cấp (Hierarchical Model) **dạng cây phân cấp**.

Trong đó:

- Mỗi nút của cây biểu diễn một thực thể
- Giữa nút con với nút cha được liên hệ với nhau theo một mối quan hệ xác định.

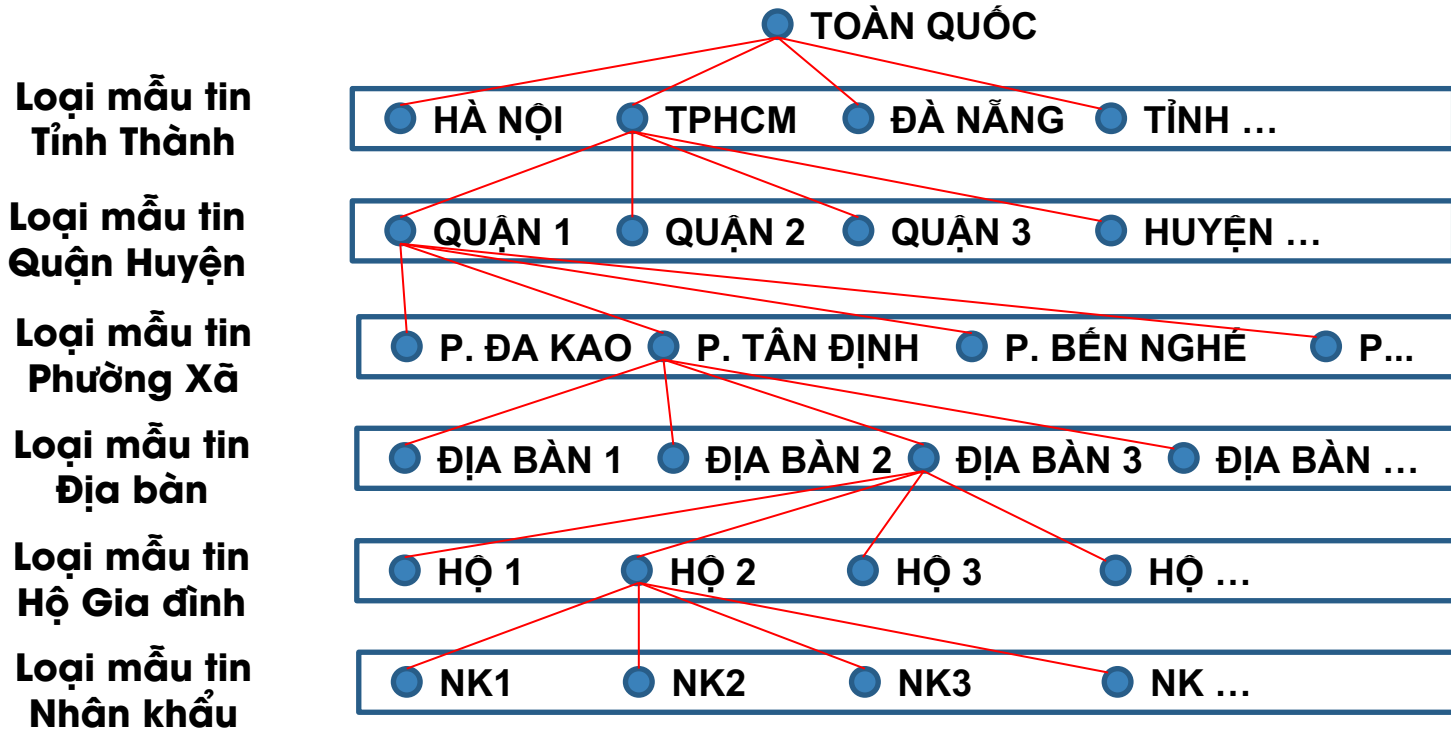


3. Mô hình dữ liệu phân cấp

- Một số khái niệm:
 - Loại mẫu tin: giống mô hình dữ liệu mạng
 - Loại mối liên hệ: phân cấp
- Giữa hai loại mẫu tin **chỉ tồn tại một mối quan hệ duy nhất**. Như vậy:
 - Mối liên hệ từ mẫu tin chủ tới các mẫu tin thành viên là $1 \div n$
 - Mối liên hệ từ mẫu tin thành viên với mẫu tin chủ là $1 \div 1$



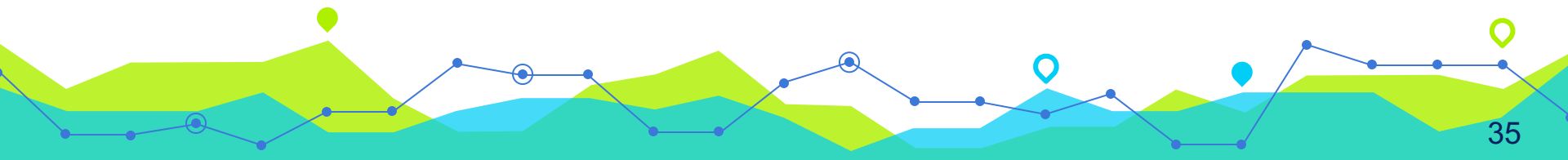
3. Mô hình dữ liệu phân cấp



3. Mô hình dữ liệu phân cấp

- Nhận xét

- Đơn giản, tìm kiếm nhanh
- Khả năng diễn đạt ngữ nghĩa kém (các mối quan hệ n-n)
- Không thích hợp biểu diễn CSDL có quy mô lớn



4. Mô hình thực thể mối kết hợp

- Mô hình thực thể mối kết hợp (ER – Entity-Relationship Model) được Peter Pin Shan CHEN giới thiệu năm 1976.
- Nhìn thế giới thực như là một tập các đối tượng căn bản được gọi là **các tập thực thể**, và **các mối quan hệ** giữa các đối tượng này.
- Mô hình ER được sử dụng nhiều trong thiết kế dữ liệu ở mức quan niệm.



4.1. Loại thực thể

- Loại thực thể (Entity Type) là những loại đối tượng hay sự vật của thế giới thực tồn tại cụ thể cần được quản lý.
- **VD:** NHANVIEN, PHONGBAN, LOP, MONHOC, SINHVIEN, ...
- Ký hiệu:

SINHVIEN

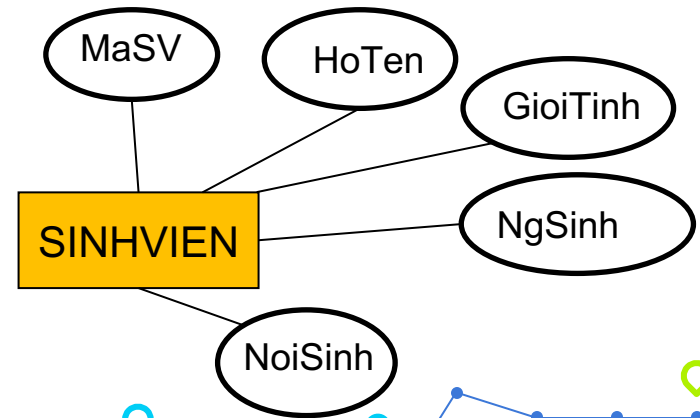
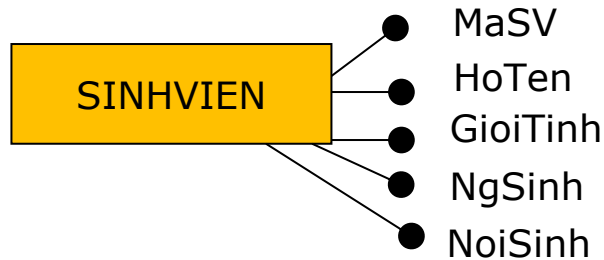
LOP

4.2. Thực thể

- Thực thể (Entity) là một thể hiện hoặc một đối tượng của một loại thực thể.
- **VD:** Loại thực thể là SINHVIEN có các thực thể:
 - ('SV001', 'Nguyễn Minh', 'Nam', '1/2/1987', 'Tp. Hồ Chí Minh')
 - ('SV002', 'Trần Năm', 'Nam', '13/2/1987', 'Đồng Nai')

4.3. Thuộc tính của loại thực thể

- Thuộc tính của loại thực thể (Entity Attribute) là những tính chất đặc trưng của loại thực thể.
- **VD:** Loại thực thể SINHVIEN có các thuộc tính:
Mã sinh viên, họ tên, giới tính, ngày sinh, nơi sinh
- Ký hiệu:



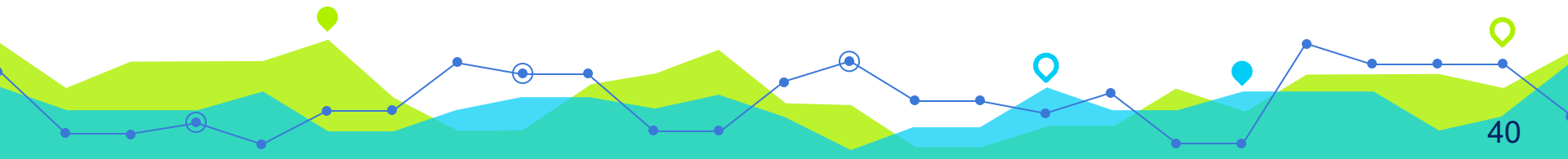
Các loại thuộc tính

- Đơn trị (Simple):

- Chỉ có một giá trị, không thể chia nhỏ.
- **VD:** MaSV, TenMonHoc, SoTC,...

- Đa hợp (Composite):

- Thuộc tính có thể được tạo thành từ nhiều thành phần.
- **VD:** DiaChi (SoNha, Duong, Phuong, Quan, Tp_Tinh)
HoTen (Ho, TenLot, Ten).



Các loại thuộc tính

- Đa trị (Multi-valued)

- Thuộc tính có thể có nhiều giá trị đối với một thực thể.
- VD:** BANGCAP ký hiệu {BANGCAP}

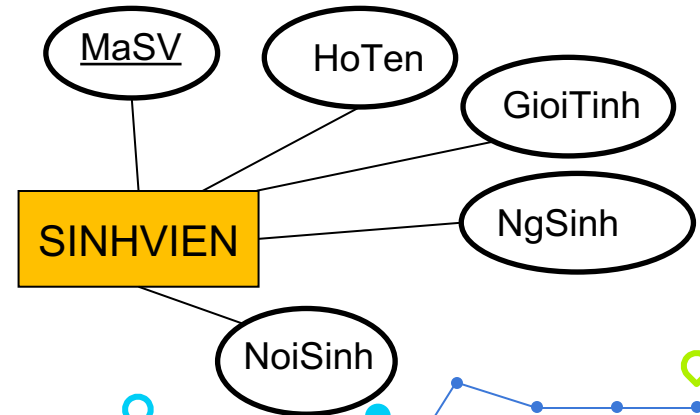
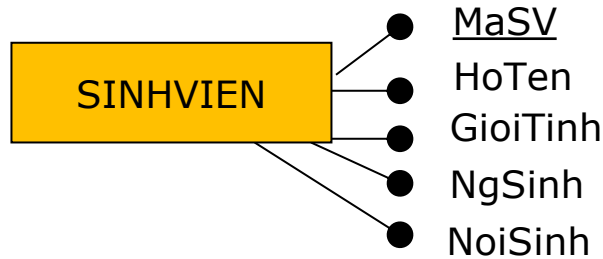
Lưu ý: Các thuộc tính đa hợp và đa trị có thể lồng nhau.

VD: BangCap của SINHVIEN có thể đa hợp và đa trị

Ký hiệu: {BangCap (NoiCap, NamCap, KetQua, ChuyenNganh)}

4.4. Khóa của loại thực thể

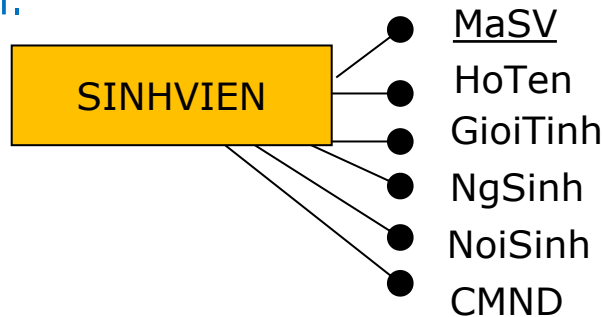
- Khóa của loại thực thể là **thuộc tính nhận diện thực thể**.
- Căn cứ vào giá trị của khóa có thể xác định **duy nhất** một thực thể.
- **VD:** Mỗi sinh viên có một mã số duy nhất → Khoá của loại thực thể **SINHVIEN** là **MaSV**
- Ký hiệu:



4.4. Khóa của loại thực thể

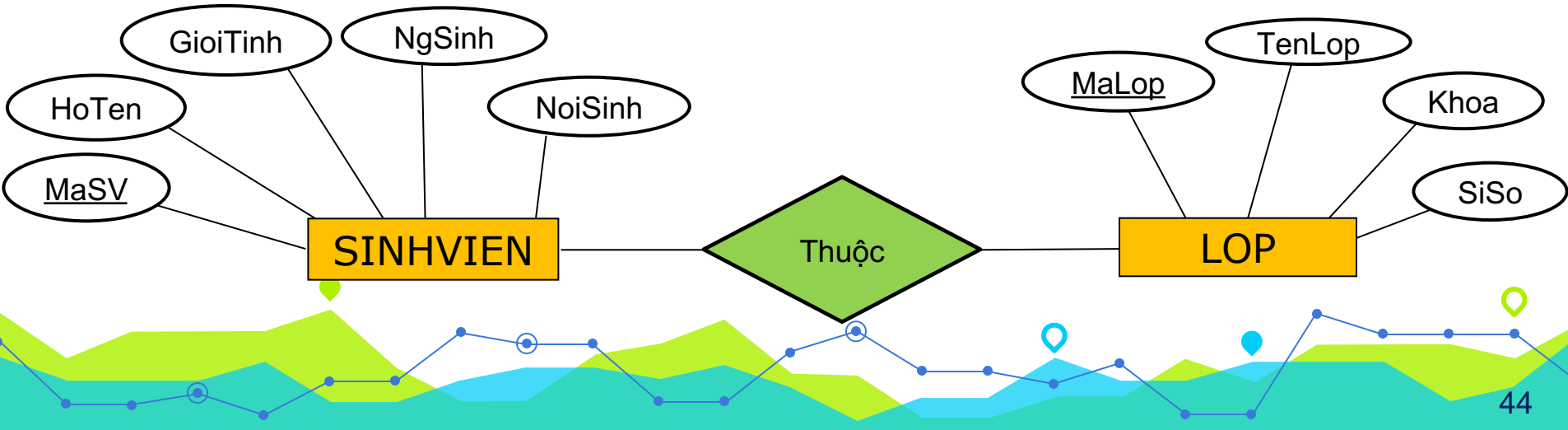
- *Lưu ý:*

- Mỗi loại thực thể có ít nhất 01 khóa.
- Một khóa có thể có 01 hoặc nhiều thuộc tính.
- Trong một loại thực thể có thể có nhiều khóa, chọn ra 01 khóa để làm khóa chính.



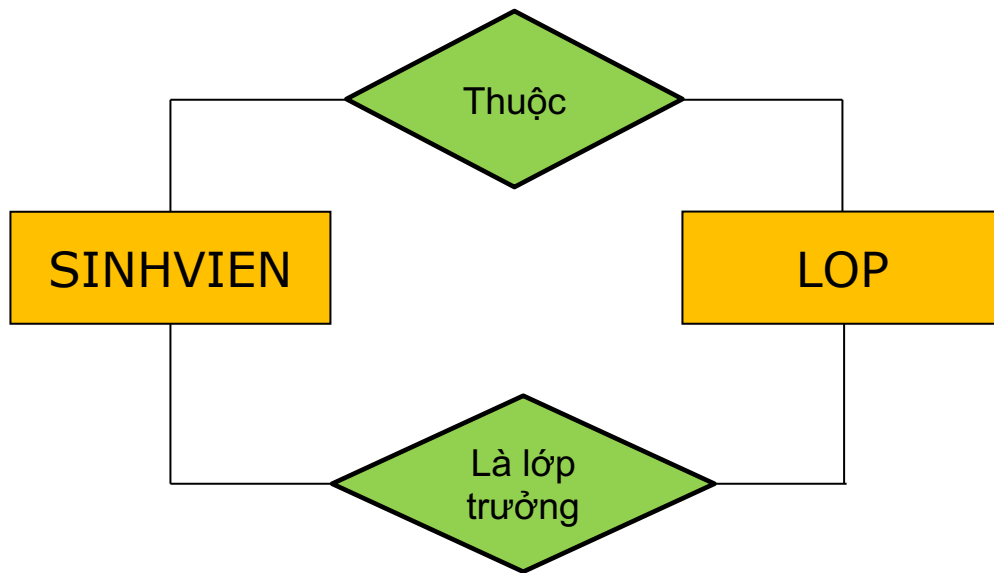
4.5. Loại mối kết hợp

- Loại mối kết hợp (Relationship type) là sự liên kết giữa hai hay nhiều loại thực thể
- Ký hiệu: bằng một hình thoi hoặc hình oval
- **VD:** Giữa SINHVIEN và LOP có loại mối kết hợp THUOC



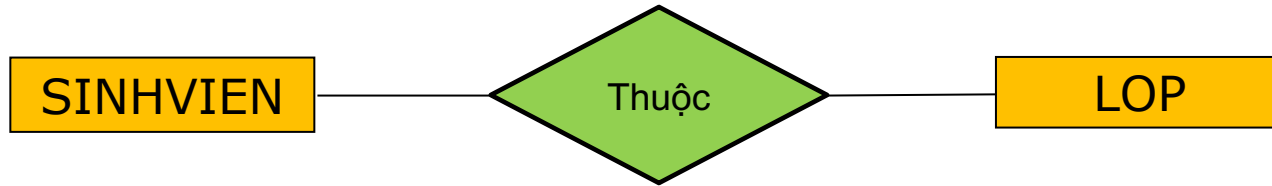
4.5. Loại mối kết hợp

- Giữa hai loại thực thể có thể tồn tại **nhiều** hơn **một** loại mối kết hợp.



4.5.1. Số ngôi của loại mỗi kết hợp

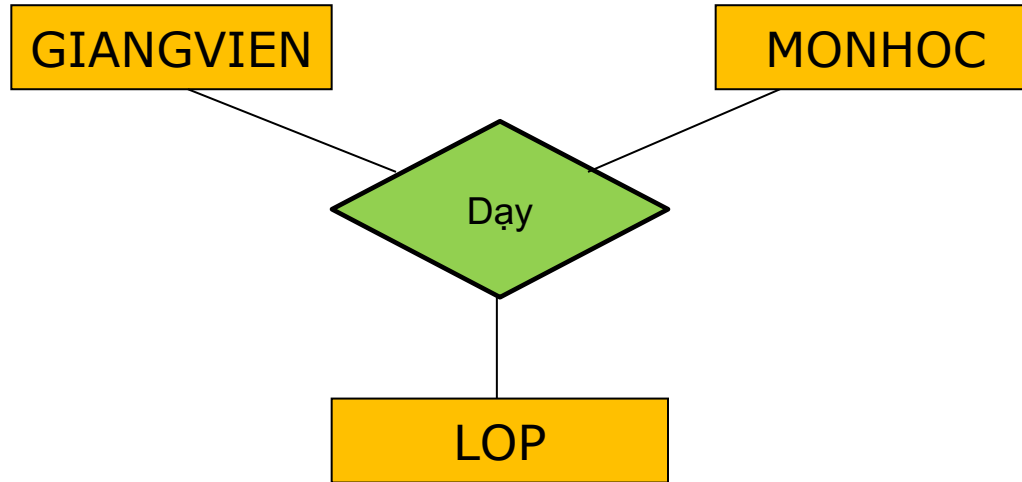
- Số ngôi của loại mỗi kết hợp (Relationship degree) là số loại thực thể tham gia vào loại mỗi kết hợp đó
- **VD:**



→ Số ngôi của loại mỗi kết hợp Thuộc: 2

4.5.1. Số ngôi của loại mỗi kết hợp

- **VD:**



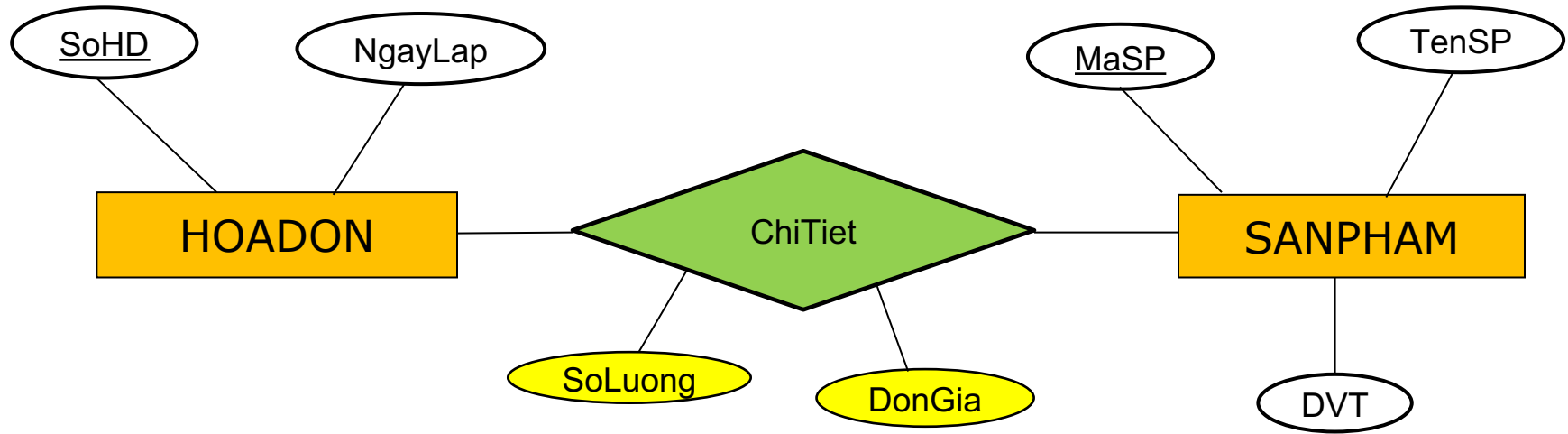
→ Số ngôi của loại mỗi kết hợp Dạy: 3

4.5.2. Thuộc tính của loại mối kết hợp

- Thuộc tính của loại mối kết hợp (Relationship type attribute):
 - Mô tả **tính chất của mối kết hợp** đó.
 - Bao gồm **các thuộc tính khóa của các loại thực thể** tham gia vào loại mối kết hợp đó, và còn có thể có thêm **những thuộc tính bổ sung** khác.
 - Không thể gắn với các thực thể tham gia vào mối kết hợp đó

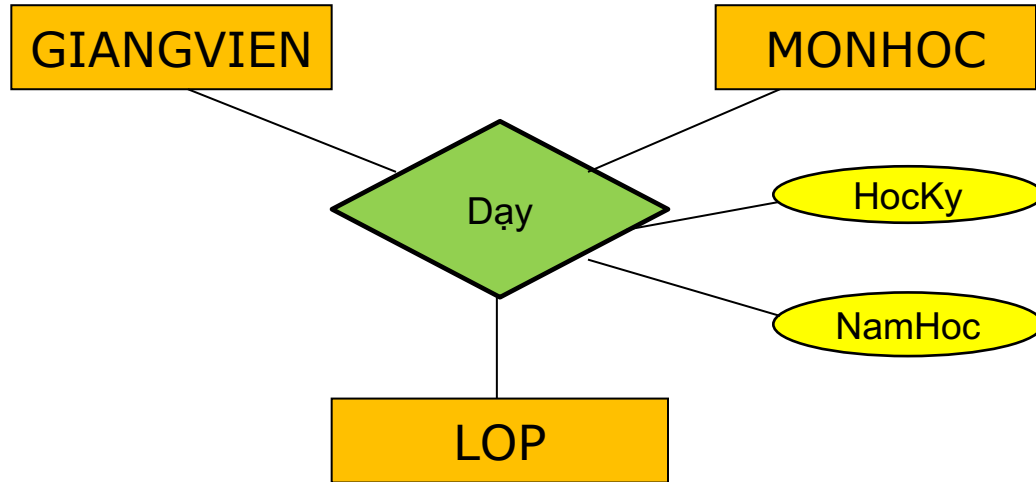
4.5.2. Thuộc tính của loại mối kết hợp

- **VD:**



4.5.2. Thuộc tính của loại mối kết hợp

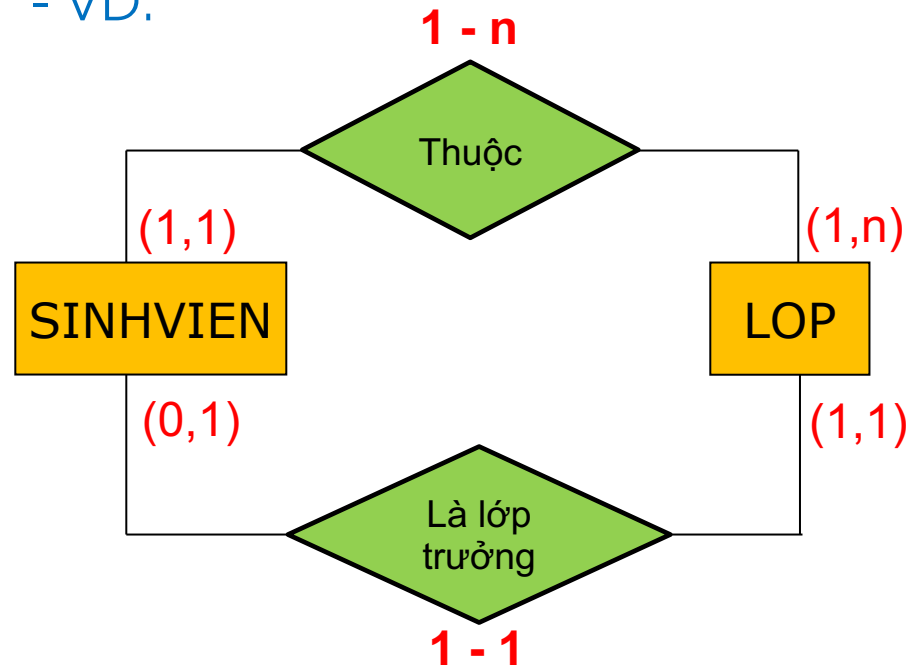
- **VD:**



4.6. Bản số

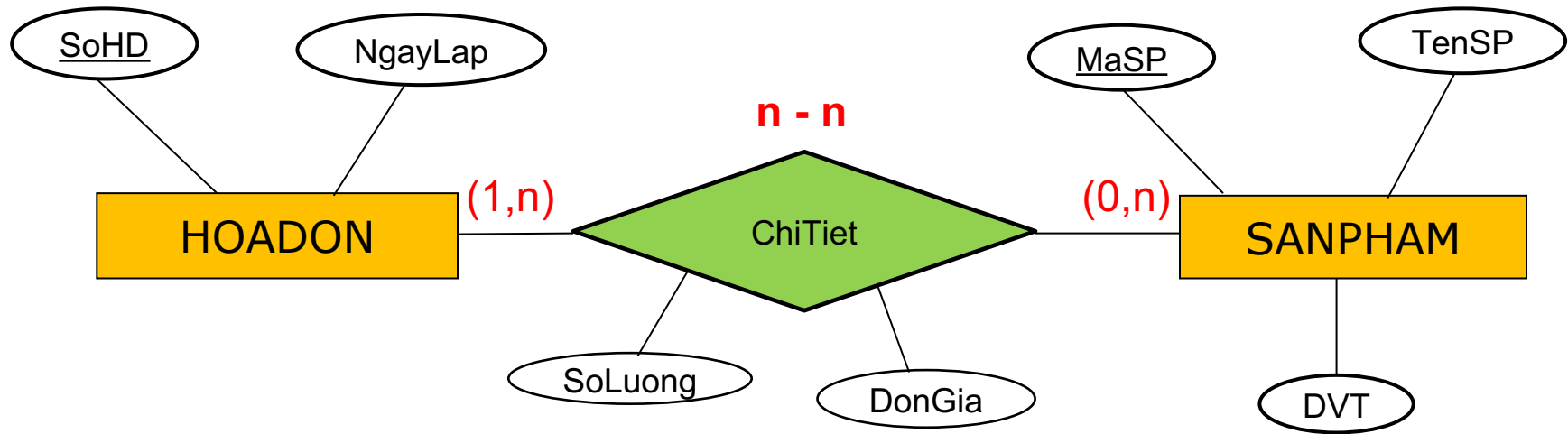
- Bản số (relationship cardinality): thể hiện số lượng tối thiểu và số lượng tối đa các thực thể tham gia vào loại mối kết hợp.
- Ký hiệu: (số lượng tối thiểu, số lượng tối đa)

- VD:



4.6. Bản số

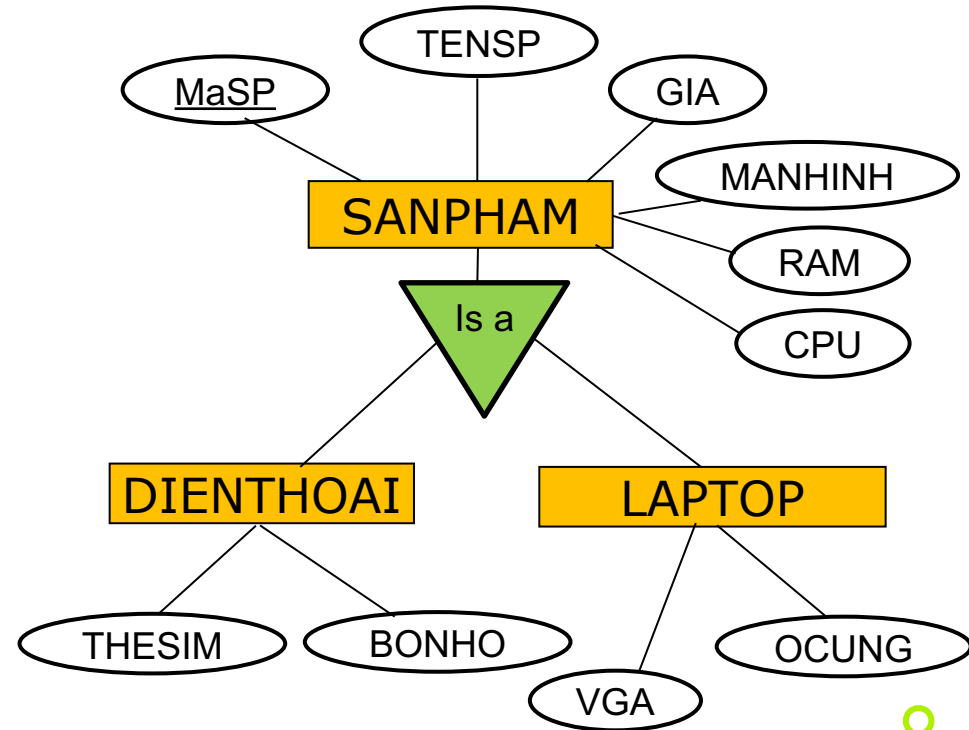
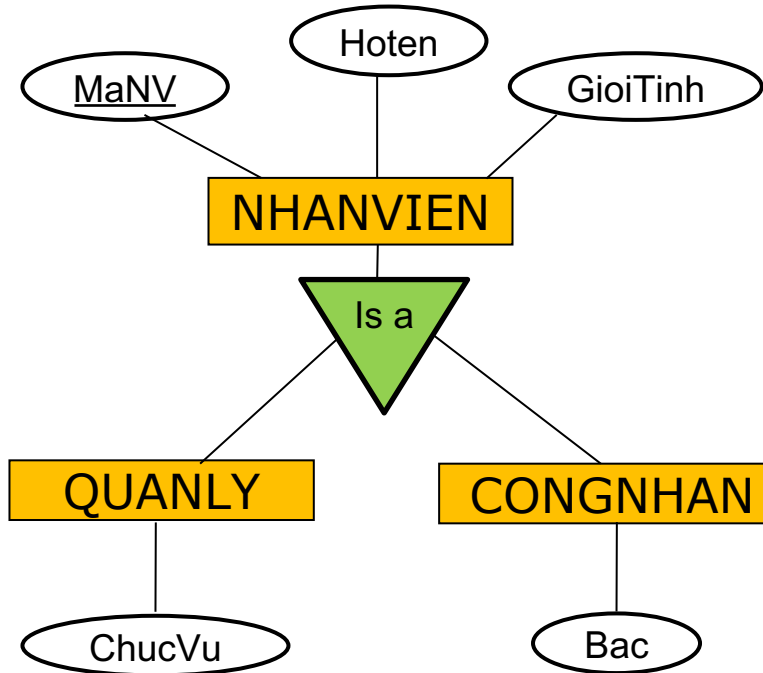
- **VD:**



4.7. Mô hình ER mở rộng

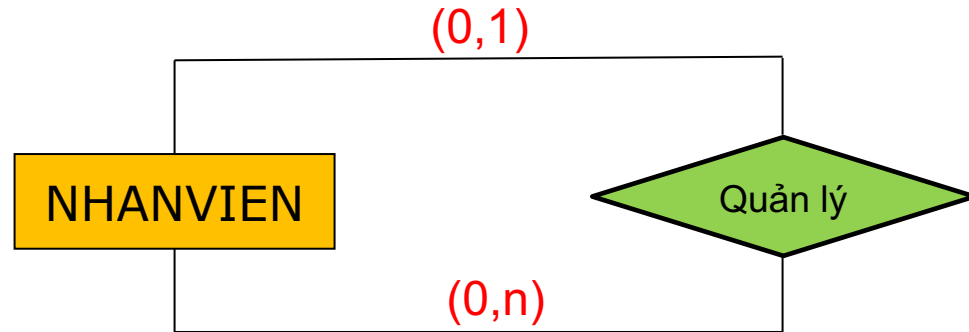
- Chuyên biệt hoá, tổng quát hóa
- Mối kết hợp đệ quy
- Loại thực thể yếu
- Mối kết hợp mở rộng

4.7.1. Chuyên biệt hóa, tổng quát hóa



4.7.2. Mỗi kết hợp đệ quy

- Là loại mối kết hợp được tạo thành từ cùng một loại thực thể (hay một loại thực thể có loại mối kết hợp với chính nó)
- **VD:** Mỗi nhân viên có một người quản lý trực tiếp và người quản lý đó cũng là một nhân viên.



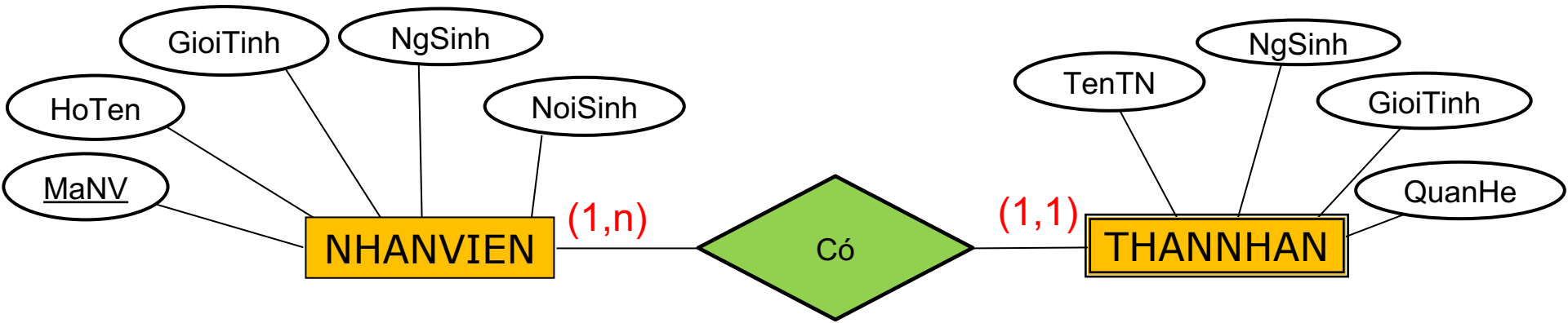
4.7.3. Loại thực thể yếu

- Loại thực thể yếu:
 - Là loại thực thể không có khóa.
 - Phải tham gia trong một loại mối kết hợp xác định trong đó có một loại thực thể chủ.
- Ký hiệu:

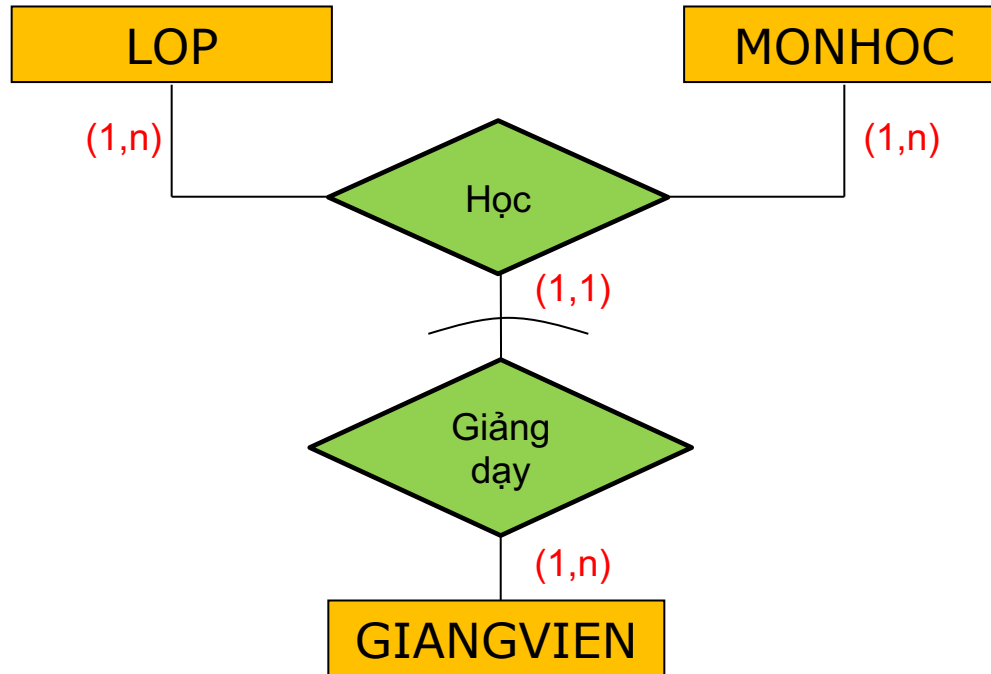
Tên Loại
thực thể yếu

4.7.3. Loại thực thể yếu

- **VD:**



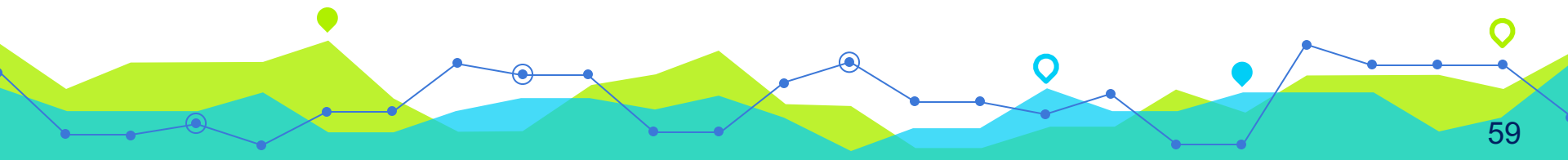
4.7.4. Mỗi kết hợp mở rộng



4. Mô hình thực thể mối kết hợp

- Các bước thiết kế:

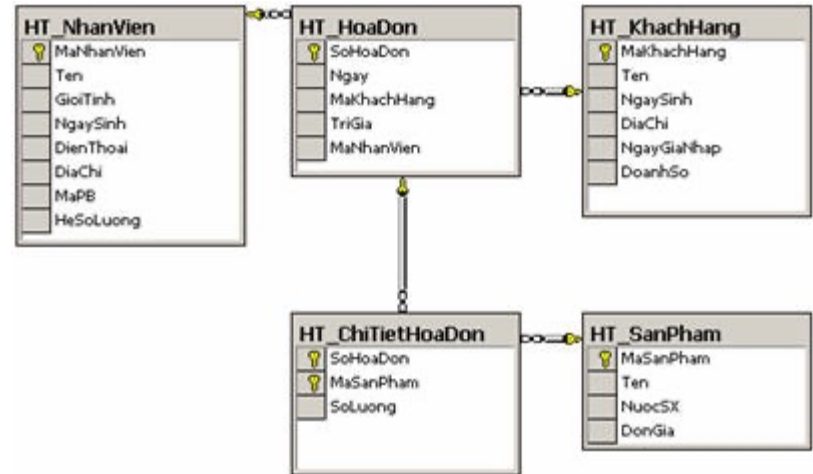
- Xác định loại thực thể
- Xác định loại mối kết hợp
- Xác định Thuộc tính của loại thực thể, loại mối kết hợp
- Xác định bản số (1-1, 1-n, n-n)



5. Mô hình dữ liệu quan hệ

VD:

- **NHANVIEN** (MaNV, HoTen, NgaySinh, DiaChi)
- **KHACHHANG** (MaKH, HoTen, NgaySinh, DiaChi, DoanhSo)
- **SANPHAM** (MaSP, TenSP, NuocSX, DonGia)
- **HOADON** (SoHD, NgayHD, MaKH, TriGia, MaNV)
- **CTHD** (SoHD, MaSP, SL)



6. Mô hình dữ liệu hướng đối tượng

- Mô hình dữ liệu hướng đối tượng (Object Oriented Data Model) ra đời vào cuối những năm 80 và đầu những năm 90
- Dựa trên **cách tiếp cận hướng đối tượng**.
- Các khái niệm: Lớp (class), Thuộc tính (Attribute), Phương thức (Operator), ...
- Đặc trưng
 - Tính kế thừa (inheritance)
 - Tính đóng gói (encapsulation)
 - Tính đa hình (polymorphism)
 - Tính tái sử dụng (reusability)

Tổng kết chương



Các khái niệm của một hệ CSDL

1. Dữ liệu (Data)
2. Hệ thống tập tin (File System)
3. Định nghĩa một CSDL
4. Các đối tượng sử dụng CSDL
5. Hệ quản trị CSDL
6. Các mức của một CSDL



Các mô hình dữ liệu

1. Mô hình dữ liệu
2. Mô hình dữ liệu mạng
3. Mô hình dữ liệu phân cấp
4. Mô hình thực thể mối kết hợp
5. Mô hình dữ liệu quan hệ
6. Mô hình dữ liệu hướng đối tượng



THANKS!

Any questions?

