>>

Mapping ER to SQL

^ >>

❖ Mapping ER to SQL

We have explored mapping ER designs to relational schemas

SQL schemas are essentially more detailed versions of relational schemas

The mapping is much the same, except that

- you need to provide more details on allowed values
- you can map some ideas from ER that are not in relational schemas

There are also some ideas from ER than do not map to an SQL schema

<< ^ >>

❖ Reminder: SQL/Relational Model vs ER Model

Correspondences between SQL/relational and ER data models:

- attribute(ER) ≅ attribute(Rel), entity(ER) ≅ row/tuple(Rel)
- entity set(ER) ≅ table/relation(Rel), relationship(ER) ≅ table/relation(Rel)

Differences between SQL and ER models:

- SQL uses tables to model entities *and* relationships
- SQL has no composite or multi-valued attributes (only atomic)
- SQL has no object-oriented notions (e.g. subclasses, inheritance)

Note that ...

- not all aspects of ER can be represented exactly in an SQL schema
- some aspects of SQL schemas (e.g. domains) do not appear in ER

<< ^ >>

❖ Mapping ER to SQL

Some conventions that we use in mapping ER to SQL

- stop using upper-case for SQL keywords (use `table` vs `TABLE`)
- all tables based on entities are given plural names
- attributes in entities are given the same name in ER and SQL
- attributes in relationships are given the same name in ER and SQL
- ER key attributes are defined using **`primary key`**
- text-based attributes are defined with type **`text`**, unless there is a size which is obvious from the context
- attribute domains can be PostgreSQL-specific types where useful
- foreign keys within entity tables are named after the relationship
- foreign keys in relationship tables are named ***`table`*`_id`**

<< ^ >>

❖ Mapping Strong Entities

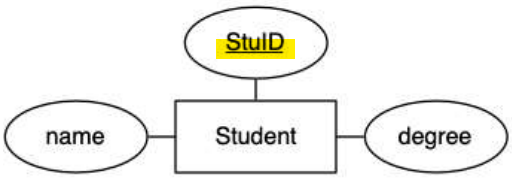
An entity set *E* with atomic attributes $a_1, a_2, \ldots a_n$

maps to

A table *R* with attributes (columns) $a_1, a_2, \ldots a_n$

Example:

ER Model



SQL Version

```
create table Students (
    stuID    integer primary key,
    name     text not null,
    degree   char(4)
);
```

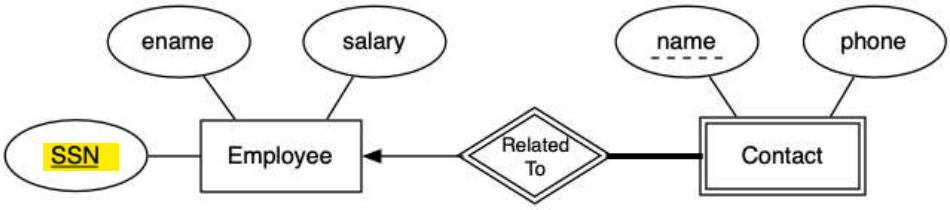
Note: the key is preserved in the mapping.

<< ^ >>

❖ Mapping Weak Entities

Example:

ER Model



SQL Version

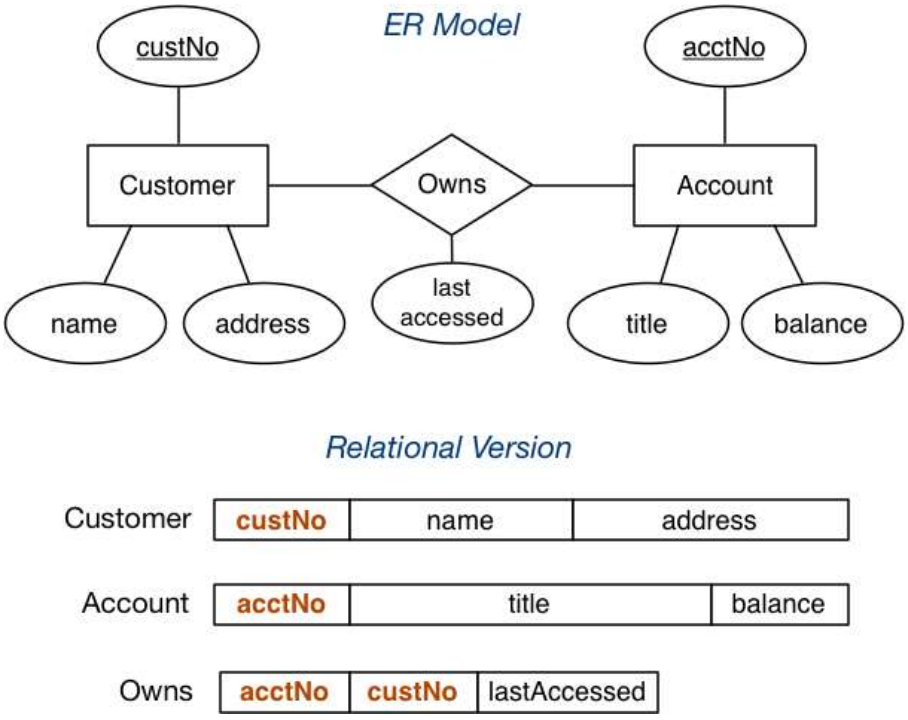
```
create table Employees (
    SSN      text primary key,
    ename    text,
    salary   currency
);
```

```
create table Contacts (
    relatedTo  text not null,  -- total participation
    name       text,           -- not null implied by PK
    phone      text not null,
    primary key (relatedTo, name),
    foreign key (relatedTo) references Employees (ssn)
);
```

<< ^ >>

❖ Mapping N:M Relationships

Example:



<< ^ >>

❖ Mapping N:M Relationships (cont)

```
create table Customers (
    custNo  serial primary key,
    name    text not null,
    address text  -- don't need to know customer's address
);
create table Accounts (
    acctNo   char(5) check (acctNo ~ '[A-Z]-[0-9]{3}'),
    title    text not null,     -- acctNos are like 'A-123'
    balance  float default 0.0,
    primary key (acctNo)
);
create table Owns (
    customer_id integer references Customers(custNo),
    account_id  char(5) references Accounts(acctNo),
    last_accessed timestamp,
    primary key (customer_id, account_id)
);
```

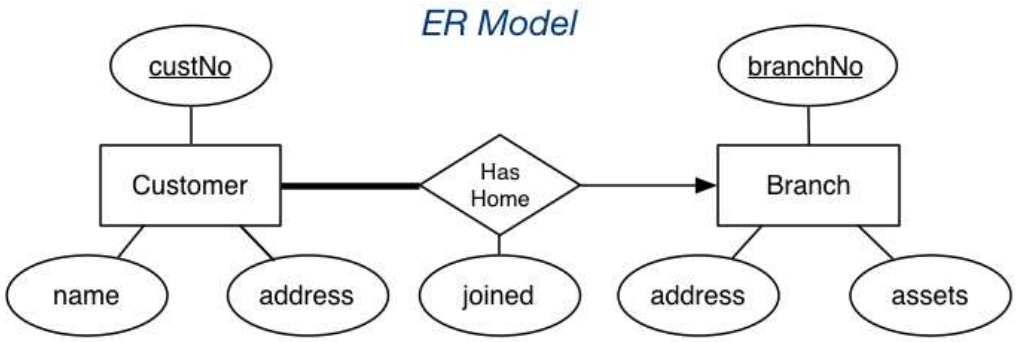
(can be null)

As it is many to many, the relationship is only uniquely defined by both the customer and the account. Hence we cannot store the relationship information on customer or account, but must instead store it on a relationship table, Owns.

<< ^ >>

❖ Mapping 1:N Relationships

Example:



Relational Version

Customer | **custNo** | name | address | branchNo | joined

Branch | **branchNo** | address | assets

<< ^ >>

❖ Mapping 1:N Relationships (cont)

```
create table Branches (
    branchNo serial primary key,
    address  text not null,
    assets   currency
);
create table Customers (
    custNo  serial primary key,
    name    text not null,
    address text,
    hasHome integer not null, -- total participation
    joined  date not null,
    foreign key (hasHome) references Branches(branchNo)
);
```

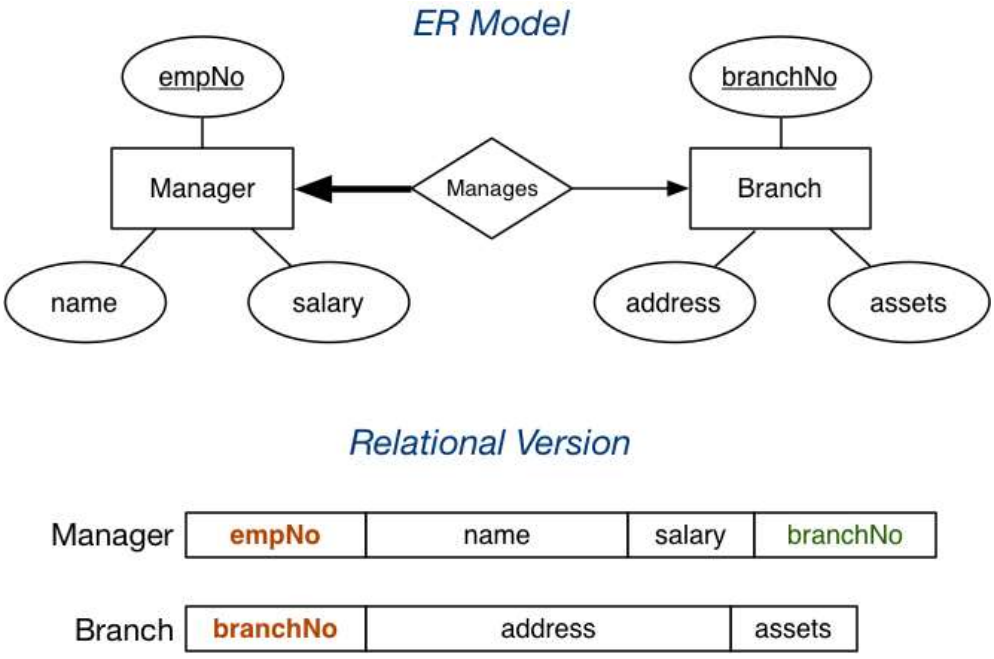
hasHome implements the 1:n relationship; **not null** implements total participation

hasHome is the branch number. It being NOT NULL means it must participate. It being a member of Customers means there can only be 1 (implements the arrow/exactly 1 behaviour).

<< ^ >>

❖ Mapping 1:1 Relationships

Example:



<< ^ >>

❖ Mapping 1:1 Relationships (cont)

```
create table Branches (
    branchNo serial primary key,
    address  text not null,
    assets   currency          -- a new branch
);                             --    may have no accounts
create table Managers (
    empNo    serial primary key,
    name     text not null,
    salary   currency not null, -- when first employed,
                                --    must have a salary
    manages  integer not null,  -- total participation
    foreign key (manages) references Branches(branchNo)
);
```

If both entities have total participation, cannot express this in SQL except by putting a (redundant) **`not null`** foreign key in one table

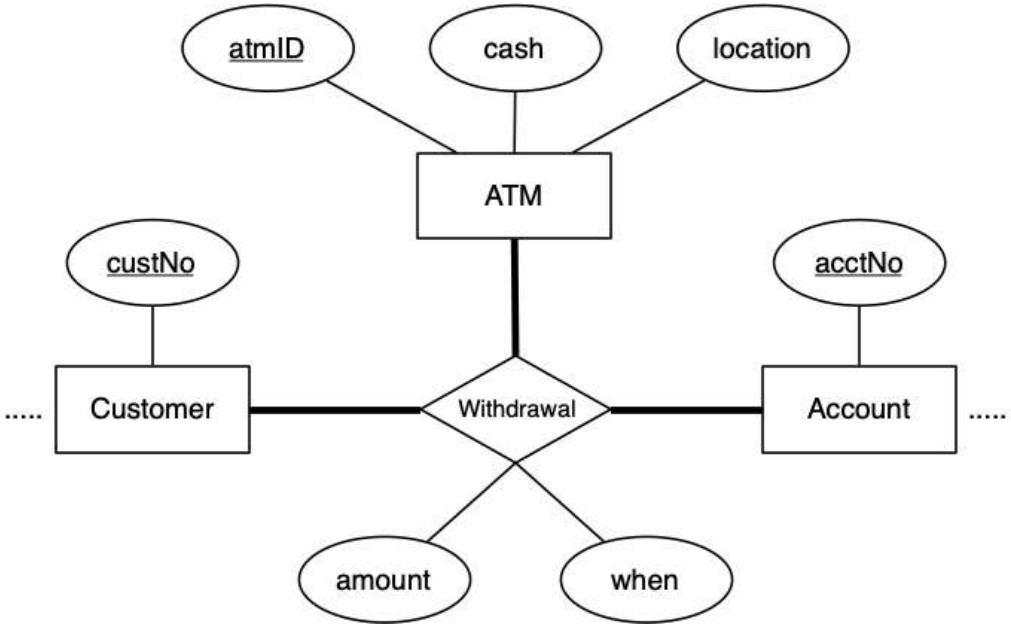
i.e. store either the branch id on the manager and a FK to manager on the branch OR the manager id on the branch, and a branch FK on the manager.

we store the branch number on the manager because the manager has total participation and therefore is guaranteed to be present in the manages relationship. Not every branch is guaranteed to be managed by a manager, hence don' t want to store the manager id here because it will be redundant for a lot of them.

<< ^ >>

❖ Mapping n-way Relationships

Example:



A customer accesses one of their accounts at a specific ATM

A customer MUST withdraw (an amount of money at a specific time) from ONE OR MORE accounts at ONE OR MORE atm's.

An atm MUST dispense an amount of money at a specific time from ONE OR MORE accounts to ONE OR MORE customers.

An Account MUST debit a certain amount of money at a specific time via ONE OR MORE atm' s to ONE OR MORE customers.

<< ^ >>

❖ Mapping n-way Relationships (cont)

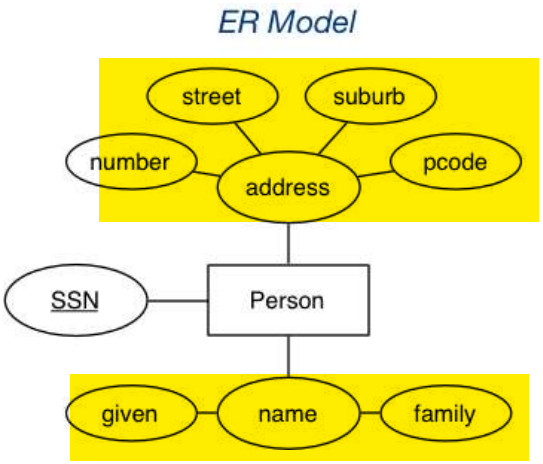
```
create table Customers (
    custNo   serial primary key, ...
);
create table Accounts (
    acctNo   char(5) ... primary key, ...
);
create table ATMs (
    atmID    serial primary key,
    cash     currency check (cash >= 0),
    location text not null
);
create table Withdrawal (
    customer_id  integer references Customers(custNo),
    account_id   char(5) references Accounts(acctNo),
    atm_id       integer references ATMs(atmID),
    amount       currency not null,
    when         timestamp default now(),
    primary key  (customer_id,account_id,atm_id)
);
```

<< ^ >>

❖ Mapping Composite Attributes

Composite attributes are mapped by concatenation or flattening.

Example:



Relational Version #1

Person

SSN	name	address

Relational Version #2

Person

SSN	given	family	.....

.....	number	street	suburb	pcode

<< ^ >>

❖ Mapping Composite Attributes (cont)

```
-- Version 1: concatenated
create table People (
    ssn     integer primary key,
    name    text not null,
    address text not null
);
-- Version 2: flattened
create table People (
    ssn     integer primary key,
    given   text not null,
    family  text,
    number  integer not null,
    street  text not null,
    suburb  text not null,
    pcode   char(4) not null check (pcode ~ '[0-9]{4}')
);
```

```
address = (number::text||' '||street||', '||suburb||' '||pcode)
```

Searching: `suburb = 'Coogee'` vs `address like '%Coogee%'`

Sorting: `order by family` vs can't be done (easily)

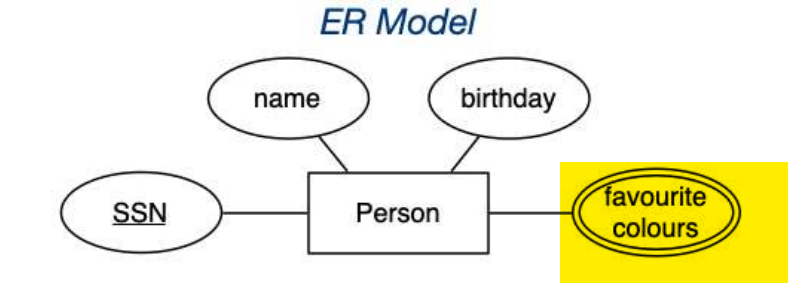
I would favour flattening. Who cares if it doesn' t look cleaner, that's not the point of a db.
Using the flattening allows for smoother operations on the stored data.
The data is also stored more rigorously because we have better constraints
i.e. we can put a char(4) not null check (pcode...) on the post code which cannot be done in the other way.

<< ^ >>

❖ Mapping Multi-valued Attributes (MVAs)

MVAs are mapped by a new table linking values to their entity.

Example:



Relational Version

Person	**SSN**	name	birthday

FavColour	**SSN**	colour

❖ Mapping Multi-valued Attributes (MVAs) (cont)

```
create table People (
    ssn      integer primary key,
    name     text not null,
    birthday date
);
create table FavColour (
    person_id integer references People(ssn),
    colour    text,
    primary key (person_id,colour)
);
```

Note that **`colour`** is implicitly **`not null`** because it is part of the primary key

<< ^ >>

❖ Mapping Subclasses

Three different approaches to mapping subclasses to tables:

- ER style
  - each entity becomes a separate table,
  - containing attributes of subclass + FK to superclass table

includes only the extension attributes and a FK to reference the superclass which has all superclass attributes (which may itself have an FK to its parent and so on).

- object-oriented
  - each entity becomes a separate table,
  - inheriting all attributes from all superclasses

Each table has its own attributes plus the attributes of all parents. This means that the further down the parent/child tree you go, the more attributes on those tables.

- single table with nulls
  - whole class hierarchy becomes one table,
  - containing all attributes of all subclasses (null, if unused)

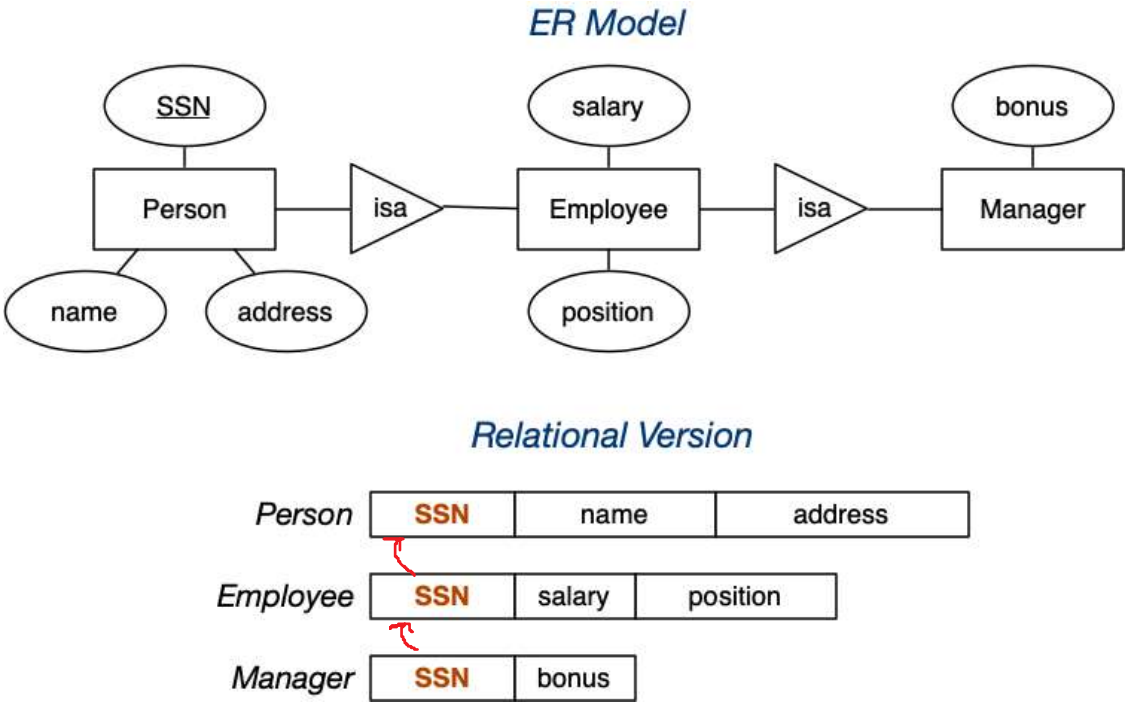
all attributes of all nodes of the tree in one table. Attributes of a node that is children to you will always be null.

Which mapping is best depends on how data is to be used.

<< ^ >>

❖ Mapping Subclasses (cont)

Example of ER-style mapping:



Here the Manager can use the SSN as a FK to look up the Employee to find the more general salary and position attributes. Likewise, Employee can then use the SSN as an FK to look-up the Person table to find even more generalised information such as name and address.

<< ^ >>

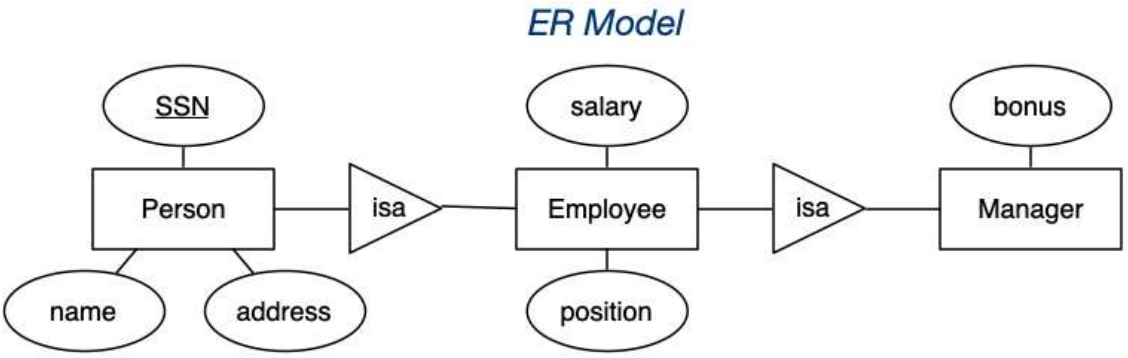
❖ Mapping Subclasses (cont)

```
create table People (
    ssn     integer primary key,
    name    text not null,
    address text
);
create table Employees (
    person_id integer primary key,
    salary    currency not null,
    position  text not null,
    foreign key (person_id) references People(ssn)
);
create table Managers (
    employee_id integer primary key,
    bonus       currency,
    foreign key (employee_id)
                references Employees(person_id)
);
```

<< ^ >>

❖ Mapping Subclasses (cont)

Example of object-oriented mapping:



Relational Version

Person	SSN	name	address			
Employee	SSN	name	address	salary	position	
Manager	SSN	name	address	salary	position	bonus

<< ^ >>

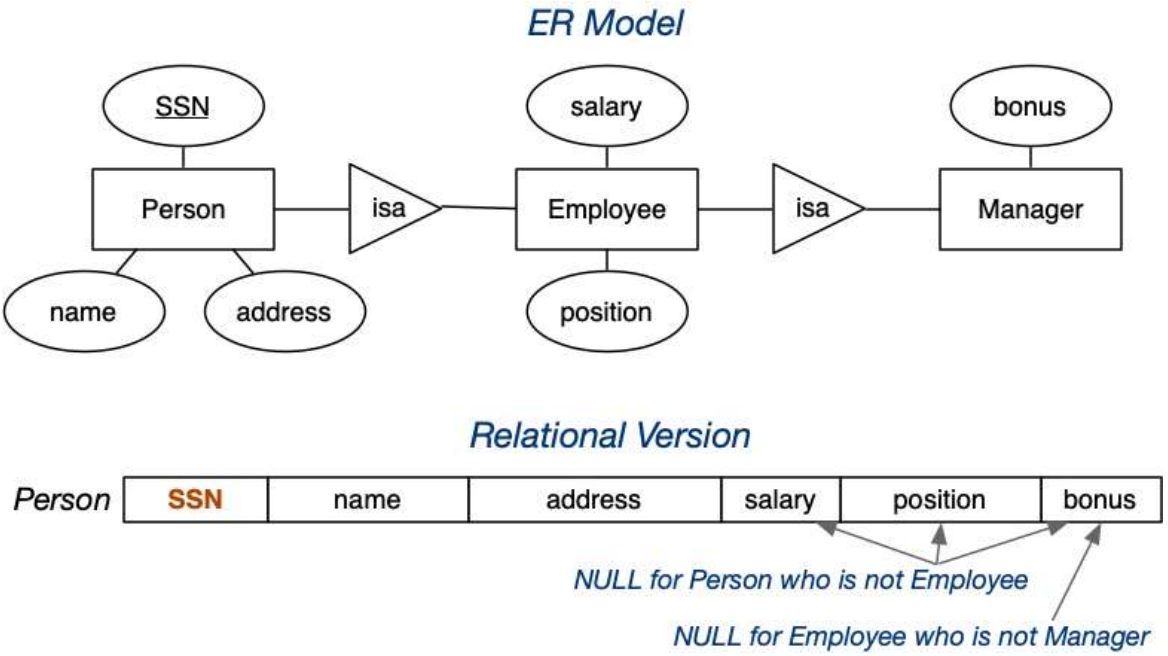
❖ Mapping Subclasses (cont)

```
create table People (
    ssn      integer primary key,
    name     text not null,
    address  text
);
create table Employees (
    ssn      integer primary key,
    name     text not null,
    address  text
    salary   currency not null,
    position text not null,
    foreign key (snn) references People(ssn)
);
create table Managers (
    ssn      integer primary key,
    name     text not null,
    address  text
    salary   currency not null,
    position text not null,
    bonus    currency,
    foreign key (snn) references People(ssn)
);
```

<< ^ >>

❖ Mapping Subclasses (cont)

Example of single-table-with-nulls mapping:



<< ^ >>

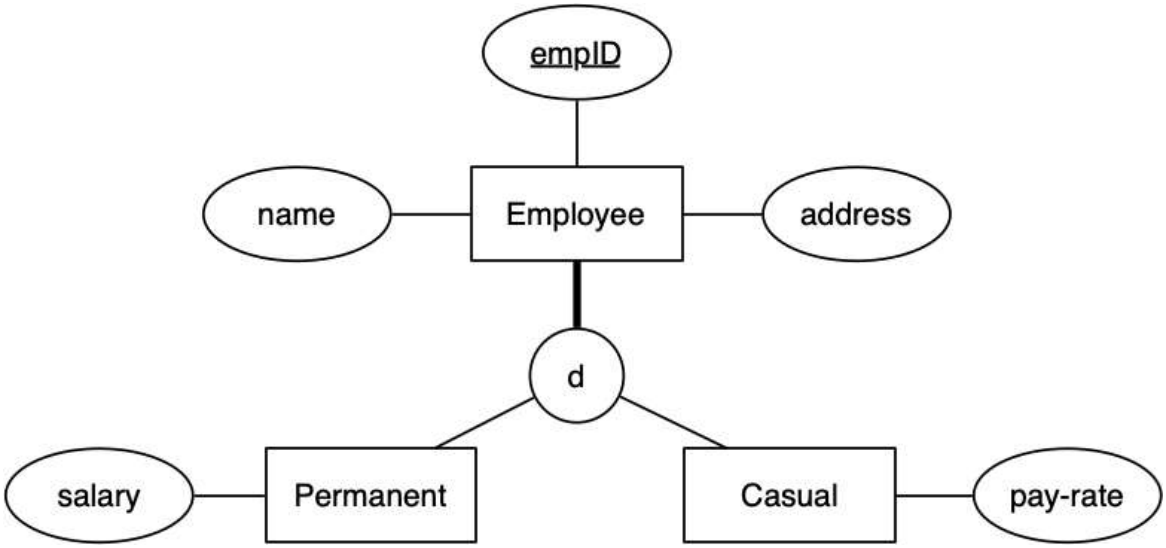
❖ Mapping Subclasses (cont)

```
create table People (
    ssn       integer primary key,
    ptype     char(1)  not null
                       check (ptype in ('P','E','M')),
    name      text not null,
    address   text
    salary    currency,
    position  text,
    bonus     currency,
    constraint subclasses check
                ((ptype = 'P' and salary is null
                 and position is null and bonus is null)
                or
                 (ptype = 'E' and salary is not null
                  and position is not null and bonus is null)
                or
                 (ptype = 'M' and salary is not null
                  and position is not null and bonus is not null))
);
```

<< ^ >>

❖ Mapping Subclasses (cont)

Example:



Every employee is either permanent or casual, but not both.

<< ^

❖ Mapping Subclasses (cont)

ER-style mapping to SQL schema:

```
create table Employees (
    empID   serial primary key,
    name    text not null,
    address text not null
);
create table Permanents (
    employee_id integer primary key,
    salary      currency not null,
    foreign key (employee_id) references Employees(empID)
);
create table Casuals (
    employee_id integer primary key,
    pay_rate    currency not null,
    foreign key (employee_id) references Employees(empID)
);
```

Does *not* capture either participation or disjoint-ness constraints!

Would need to program a solution to this e.g web-form that requires user to enter both Employee and subclass info

i.e. there is nothing stopping us currently from adding an employee to both the permanents and casuals tables. There is also nothing enforcing us to add either one for an employee i.e. it is currently possible for an employee to be neither casual nor permanent which is forbidden by the ER.

Produced: 22 Sep 2020