Medical Informatics

Lecture 2: Database Design

Dr Areti Manataki

Nanjing Medical University

Introduction to databases

- Database: a well-organised collection of data that are related in a meaningful way, describing a domain of interest
  - Examples: university, library, biomedical

Introduction to databases

- Relational databases: focus on *things* in the domain of interest and the *relations* between them
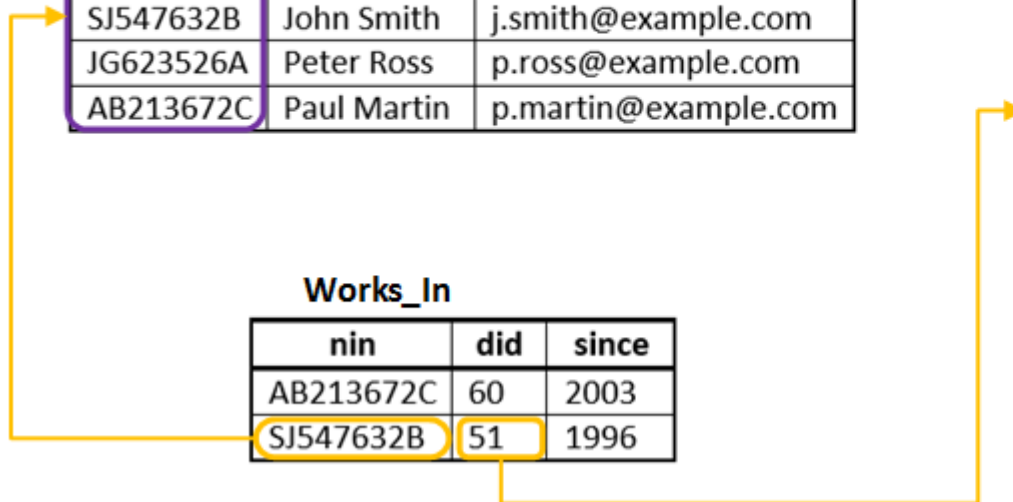
Employee

nin	name	email
SK728468L	Kate Taylor	k.taylor@example.com
SJ547632B	John Smith	j.smith@example.com
JG623526A	Peter Ross	p.ross@example.com
AB213672C	Paul Martin	p.martin@example.com

Department

did	dname	budget
51	Information Technology	80,000
56	Human Resources	50,000
60	Accounting	40,000

Works_In

nin	did	since
AB213672C	60	2003
SJ547632B	51	1996

Introduction to databases

- Database management system: software that helps maintain and utilise large collections of data in an efficient manner
  - Examples: Ms Access, Ms SQL Server, Oracle DBMS, MySQL and many more
- Advantages include:
  - Data independence
  - Efficient data access
  - Data integrity
  - Concurrent access and recovery from crashes

In the next lectures

- Data representation
    - Entity-relationship model
    - Relational model
- Data manipulation
   - SQL

Database design

- Requirements Analysis
  - What do the users want from the database?
  - Understand what data is to be stored in the database, what applications must be built on top of it, etc.

- Conceptual Database Design
  - High-level description of the data to be stored in the database, along with constraints over the data.

- Logical Database Design
  - Implement the conceptual design by mapping it to a specific data representation.

Entity Relationship Model

- It allows us to describe the data involved in a domain of interest in terms of *entities* and the *relationships* between them.

- It provides a high-level model for conceptual database design.

- It can be easily converted to a logical data model, such as the relational model.

- It can be graphically represented through an Entity Relationship diagram, which helps clarify and communicate the conceptual design.

Entities & Entity sets

- Entity: an object in the real world that is distinguishable from other objects
  - Examples: Areti, Edinburgh, Business School, Room 151
- Entity set: a collection of similar entities
 - Examples: Employee, City, University

Employee
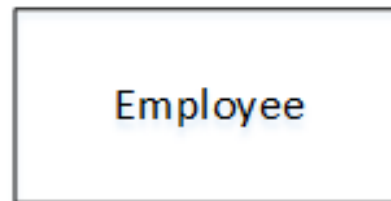
Attributes

- Each entity from the same entity set has characteristic attributes.
  - Examples: name, nin, email
- Each attribute has a domain of allowed values
  - For example, "name" has domain "string of up to 20 characters".
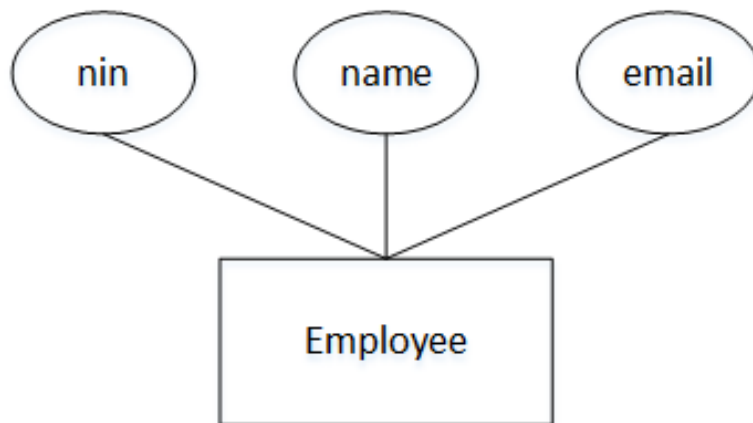
Keys

- Key: a *minimal* set of attributes whose values *uniquely identify* an entity in an entity set
- There could be more than one candidate keys. In this case, we specify one of them to be the primary key.
  - For example, both "nin" and "email" are candidate keys for Employee.
 - We designate "nin" to be the primary key.
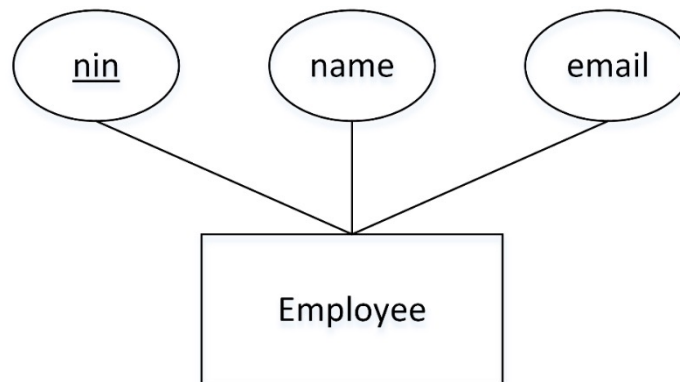
Relationships & Relationship sets

- Relationship: an association among two or more entities
  - Examples: Works_In, Lives_In, Teaches
- Relationship set: a collection of similar relationships
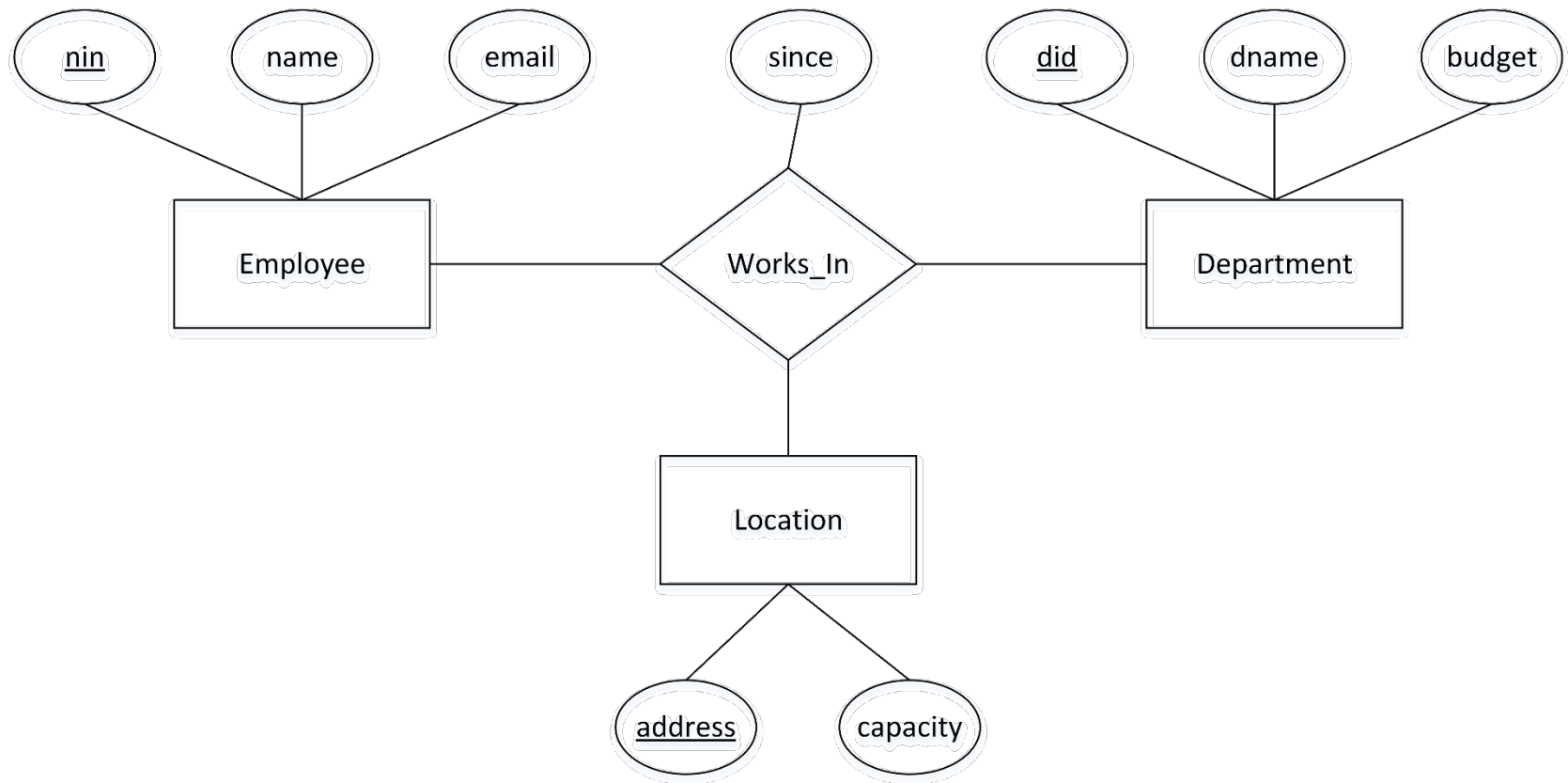- Relationships may have attributes of their own
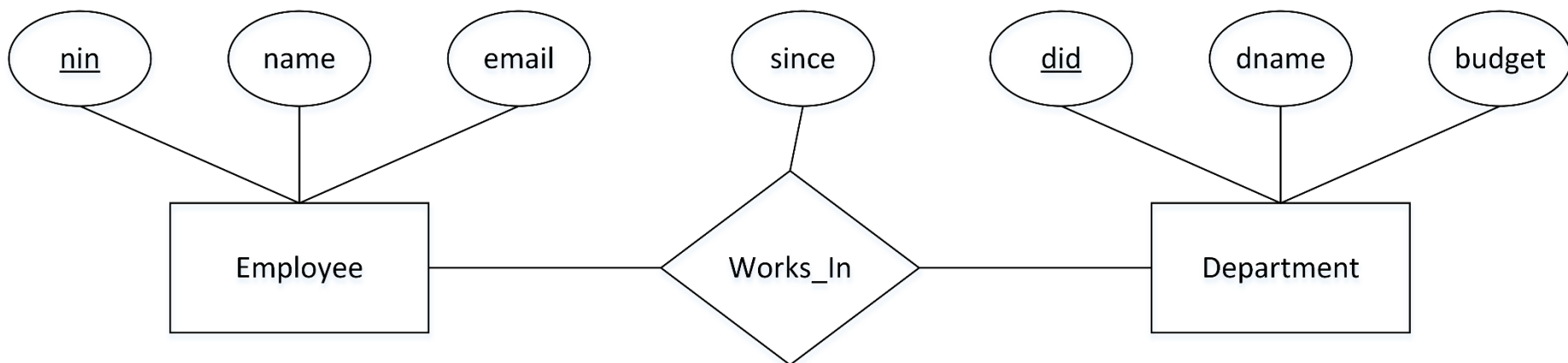
Relationships & Relationship sets

- An entity set may be involved in any number of different relationship sets
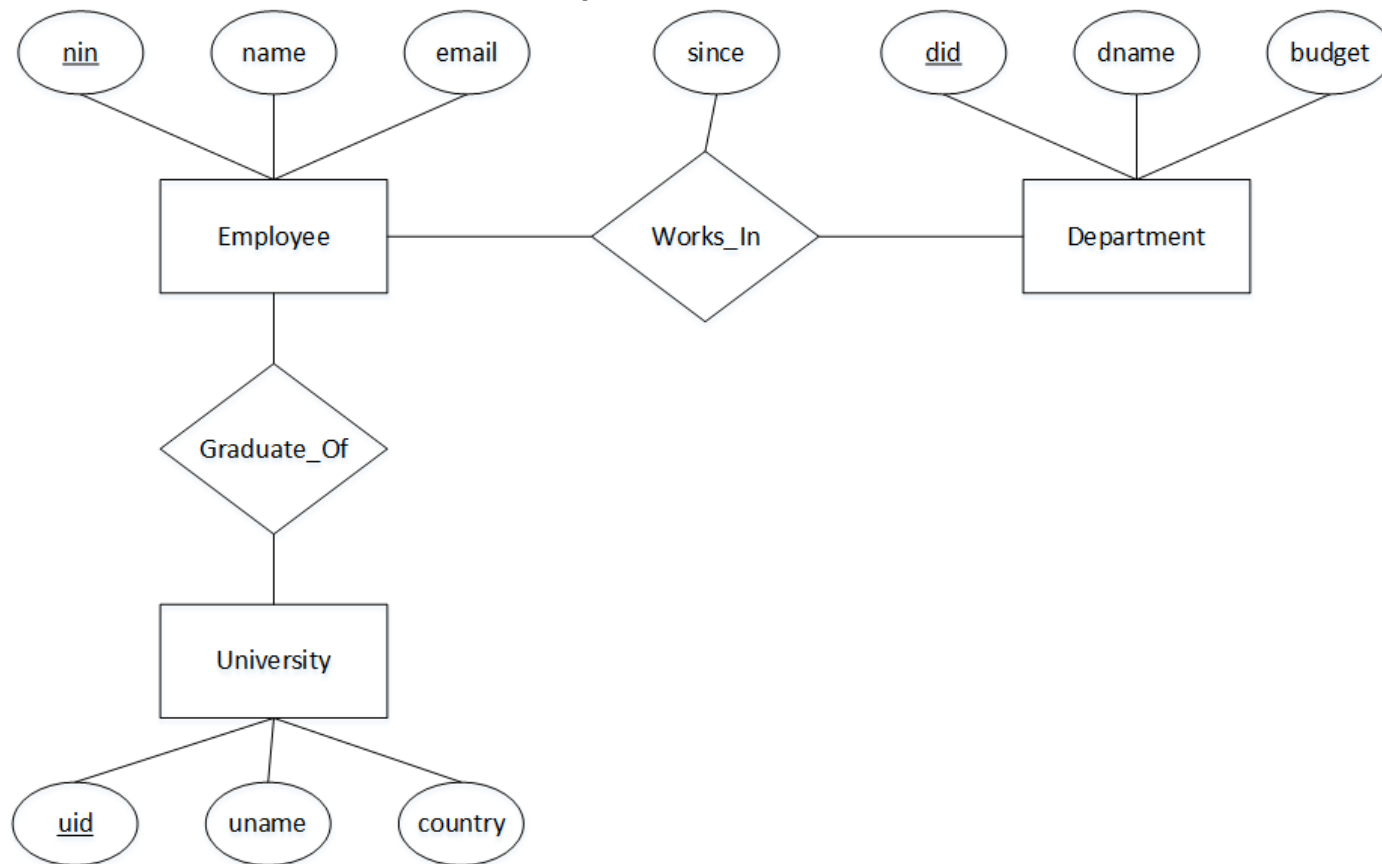
Relationships & Relationship sets

- Several entity sets may participate in the same relationship set

Instances

- Entity instance: the set of attribute values for a particular entity in an entity set
    - For example,
      (SJ547632B, John Smith, j.smith@example.com)
      is an entity instance from the "Employee" entity set.
- Relationship instance: an individual occurrence of a relationship within a relationship set, as captured by the set of associated attribute values
    - For example, (SJ547632B, 51, 1996) is a relationship instance from the "Works_In" relationship set.

Conclusions

- The ER model is used for the conceptual design of a database.

- Main constructs: entities, relationships, attributes, keys

- An ER data model can be graphically represented through an ER diagram.

- The ER model allows us to also specify restrictions on entities and the relationships they participate in. We will have a look at this in the next lecture.

Acknowledgements

The content of these slides was originally created for the Medical Informatics course from The University of Edinburgh, which is licensed under a Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0) license.

These lecture slides are also licensed under a CC BY-SA 4.0 license.