Computer Science 202

Database Systems:
Entity Relation Modelling

Objectives

- To learn what a conceptual model is and what its purpose is.
- To learn the difference between internal models and external models.
- To learn how internal and external models serve the database design process.
- To learn how relationships between entities are defined and refined, and how such relationships are incorporated into the database design process.
- To learn how ERD components affect database design and implementation.
- To learn how to interpret the modelling symbols for the four most popular E-R modelling tools.
- To learn that real-world database design often requires you to reconcile conflicting goals.

What is Modelling?

- Modelling is the process of creating a graphical representation of the database structures, and entities
- Creation of a visual representation of data which cannot otherwise be seen
- Visual models are often easier to understand

The Data Model

- Outcome of the data modelling process
- Produces a relatively simple graphical representation of the complex real-world data types within the DBMS

Degrees of Abstraction

- When creating models various levels of abstraction can be used
  - Conceptual
  - Internal
  - External
  - Physical

Conceptual

- Global view of data
- Very high-level overview often used by management
- Basis for the Identification and description of the main objects about which data will be stored
- Platform independent

Internal

- Used once DBMS software has been selected (often software specific)
- Adaptation of the conceptual model to the requirements/constraints of the DBMS
- Particularly important for Network and Hierarchical databases

External

- The view of data structures, which users and applications will see
- Makes use of table views (SQL/RDBMS)
- Often a subset of the internal data views
- Ensures Security constraints
- Aids in application development

Physical

- This is the lowest level of abstraction
- Dependant on both software and the underlying hardware
- Creation of physical storage containers (DB volumes)
- Physical layout can be very important for performance.

The Entity Relationship Model

- E-RM is used for representing the conceptual view of data
- Consists of
  - Entities
    - Corresponds to the entire table
    - Rows are instances of a physical entity
  - Attributes
    - Primary key attribute's) are indicated by being **underlined**
  - Relationships

E-RM: Attributes

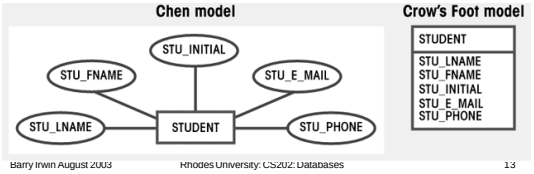
- A characteristic of an entity
- 'something that you want to know about an entity'
- Each attribute has a domain – the range of possible values which can be assigned to it

ERM: Attribute types

- Simple
  - Cannot be subdivided (e.g. Age, gender)
- Composite
  - Can be broken into further components
  - e.g. address → Box, Town, postal code
- Single valued
  - An attribute of which there can be only one (e.g. ID number)
- Multi-valued
  - May have many different values (e.g. phone number)
- Derived
  - Attribute value is derived by a formula based on other attributes

ERM Representations

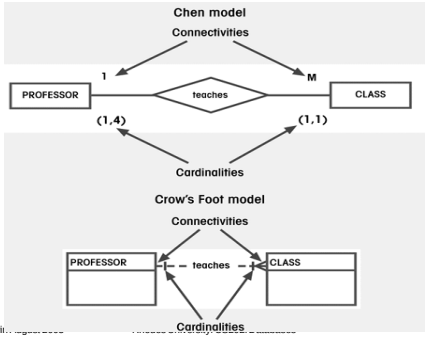
STUDENT(
STU_LNAME,STU_FNAME,STU_INITIAL
STU_E_MAIL, STU_PHONE)



Relationships

- An association between entities
- Entities are participants in a relationship
- Relation is bidirectional
- Relation is usually read from left to right
- Connectivity describes the classification
  - 1:1 , 1:M , M:N
- Cardinality
  - The number of entity occurrences associated with a single occurrence of the related entity

Cardinality and Connectivity



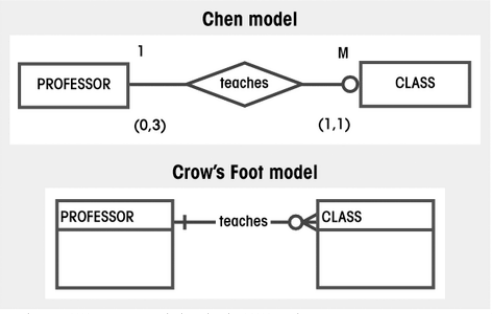
Relationship Strength

- Existence dependence
  - An Entity's existence is dependant on related entities
  - Existence-independent entities can exist apart from related entities
    - DEPENDENT has existence relation to EMPLOYEE
- Weak (non-identifying)
  - One entity is existence-independent on another
  - PK of related entity doesn't contain PK component of parent entity
- Strong (identifying)
  - One entity is existence-dependent on another
  - PK of related entity contains PK component of parent entity

Relationship Participation

- Mandatory
  - Occurrence of Entity A requires a corresponding occurrence in related Entity B
  - If a an optional symbol is omitted form an ERD relationship is mandatory
- Optional
  - Occurrence of Entity A requires a corresponding occurrence in related Entity B
  - Indicated on ERD by using a small circle on the optional side of the relationship

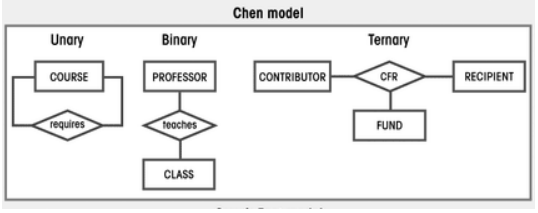
Relationship Participation



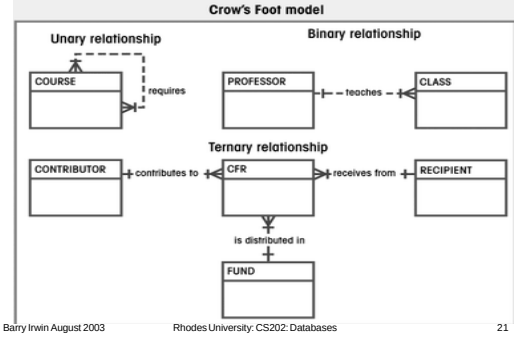
Relationship Degree

- Determined by the number of associated entities
- Unary
  - Single entity with a recursive relation to itself
- Binary
  - Two associated entities
- Ternary
  - Three associated entities

Relationship Degree: CHEN



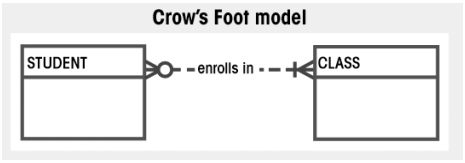
Relationship Degree: Crow's Foot



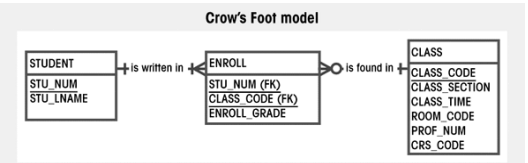
Composite Entities

- These are use for bridging between entities in an M:N relationship
- Composed at a minimum of the primary keys of both entities in the M:N relation
- Reduced the M:N relation to two 1:M relations

Composite Entities



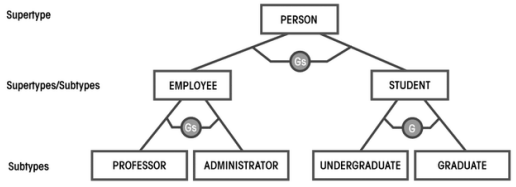
Composite Entities



Entity Super types & Subtypes

- Generalization hierarchy
  - Depicts relationships between higher-level supertype and lower-level subtype entities
  - Supertype has shared attributes
  - Subtypes have unique attributes
  - Disjoint relationships
    - Unique subtypes
    - Non-overlapping
    - Indicated with a 'G'
  - Overlapping subtypes use 'Gs' Symbol

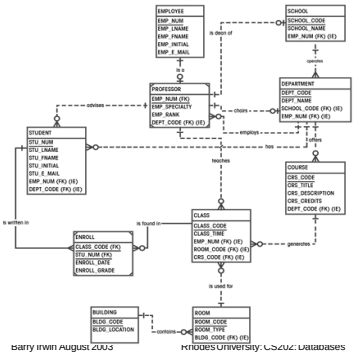
Generalization Hierarchy with Overlapping Subtypes



ERD Development

- Developing an ERD is an iterative process
  1. General outline of organisational operation developed
  2. Basic Graphical ER model created
  3. Modifications made to ER to include new entities as the diagram and relationship is refined
- Continue with refinement until DB designers and users agree on output

Completed ERD



Final outcome of the worked example
On pages R&C pages 149-157

R&C page 157
Fig 3.52

Challenge of Database Design

- Database must conform to organisation design standards
- High-speed processing may require design compromises for performance reasons
- Quick responses may be a overriding design goal
- Other concerns
  - Security
  - Performance
  - Shared access
  - Integrity