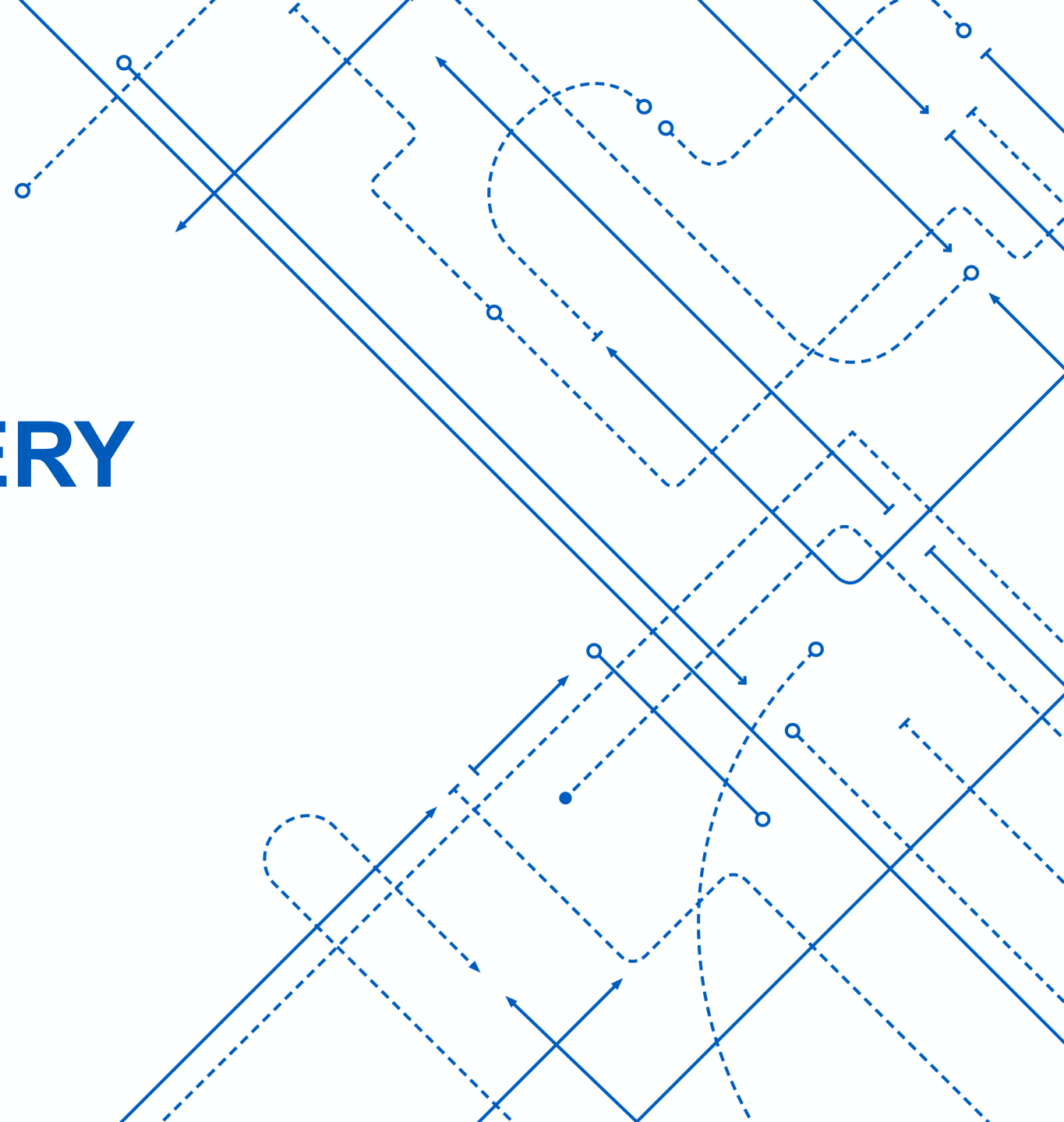


CSE460/560 DATA MODELS AND QUERY LANGUAGES

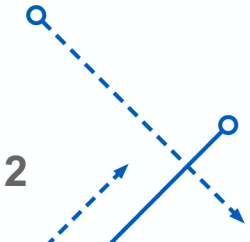
Relational Database Design

Cheng-En Chuang
(Slides Adopted from Jan Chomicki and Ning Deng)

University at Buffalo
Department of Computer Science and Engineering
School of Engineering and Applied Sciences

Outline

1. Good and Bad DB Schemas
2. Functional Dependencies
    1. Functional Dependencies
    2. Inference of FDs
3. Normal Forms
    1. BCNF and 3NF
4. Decomposition
    1. Decomposition into BCNF
    2. Decomposition into 3NF
5. Multivalued Dependencies(MVDs) and 4NF



Outline

1. **Good and Bad DB Schemas**
2. Functional Dependencies
    1. Functional Dependencies
    2. Inference of FDs
3. Normal Forms
    1. BCNF and 3NF
4. Decomposition
    1. Decomposition into BCNF
    2. Decomposition into 3NF
5. Multivalued Dependencies(MVDs) and 4NF

Good and Bad DB schemas

- Bad Schema
  - **Repetition** of information
    - Leads to **redundancies**
    - Potential inconsistencies
    - Update **anomalies**
  - Inability to represent information
    - Leads to anomalies in insertion and deletion

title	*year*	*length*	*genre*	*studioName*	*starName*
Star Wars	1977	124	SciFi	Fox	Carrie Fisher
Star Wars	1977	124	SciFi	Fox	Mark Hamill
Star Wars	1977	124	SciFi	Fox	Harrison Ford
Gone With the Wind	1939	231	drama	MGM	Vivien Leigh
Wayne's World	1992	95	comedy	Paramount	Dana Carvey
Wayne's World	1992	95	comedy	Paramount	Mike Meyers

Good and Bad DB schemas

- Good Schema
  - Relation schemas in normal form
    - Redundancy and anomaly-free
    - BCNF, 3NF
  - Schema Decomposition
    - Improving a bad schema
    - Desirable properties
      - Lossless-join
      - Dependency Preservation

Integrity Constraints

- Functional Dependencies
  - Key constraints cannot express uniqueness properties in a proper subset of all attributes
  - Key constraints need to be generalized to functional dependencies
- Other Constraints
  - Not relevant for decomposition

Outline

1. Good and Bad DB Schemas
2. **Functional Dependencies**
    1. Functional Dependencies
    2. Inference of FDs
3. Normal Forms
    1. BCNF and 3NF
4. Decomposition
    1. Decomposition into BCNF
    2. Decomposition into 3NF
5. Multivalued Dependencies(MVDs) and 4NF

Functional Dependencies (FDs)

- Notation
  - Relation schema $R(A_1, A_2, \ldots, A_n)$
  - Sets of attributes of R: $X, Y, Z, \ldots \subseteq \{A_1, \ldots, A_n\}$
  - XY instead of $X \cup Y$
- Syntax
  - $X \rightarrow Y$
    - X functionally determine Y
- Semantics
  - R satisfies $X \rightarrow Y$
    - $\forall\, t_1, t_2 \in R$
      - if $t_1[X] = t_2[X]$, *then also* $t_1[Y] = t_2[Y]$

Functional Dependencies (FDs)

- Given a relation: Movie1(title, year, length, genre, studioName, startName)
  - $title\ year \rightarrow length\ genre\ studioName$ holds
  - $title\ year \rightarrow startName$ doesn't hold
  - An example of FD doesn't hold?

title	*year*	*length*	*genre*	*studioName*	*starName*
Star Wars	1977	124	SciFi	Fox	Carrie Fisher
Star Wars	1977	124	SciFi	Fox	Mark Hamill
Star Wars	1977	124	SciFi	Fox	Harrison Ford
Gone With the Wind	1939	231	drama	MGM	Vivien Leigh
Wayne's World	1992	95	comedy	Paramount	Dana Carvey
Wayne's World	1992	95	comedy	Paramount	Mike Meyers

Recommended Reading

Database Systems: The Complete Book
Chapter 3.1