Uniqueness and Identity Rules in ORM

(Chapter 4)

Mustafa Jarrar

Birzeit University
mjarrar@birzeit.edu
www.jarrar.info

1

Watch this lecture
and download the slides

Download: http://www.jarrar.info/courses/ORM/Jarrar.LectureNotes.UniquenessANDIdentityRulesInORM.pdf
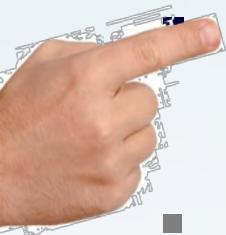
Online Courses: http://www.jarrar.info/courses/

Keywords: Identity, Uniqueness, Cardinality, Business Rules, فريد, هوية,

Uniqueness and Identity Rules in ORM

Part 1: What is Uniqueness/Identity?

- Part 2: Internal Uniqueness (within a Fact Type)

- Part 3: External Uniqueness (across Fact Types)

- Part 4: Key Length and Reference Schema

Conceptual Schema Design Steps

1. From examples to elementary facts

2. Draw fact types and apply population check

3. Combine entity types

4. Add uniqueness constraints

5. Add mandatory constraints

6. Add set, subtype, & frequency constraints

7. Final checks, & schema engineering issues

# Identity Criteria and Uniqueness
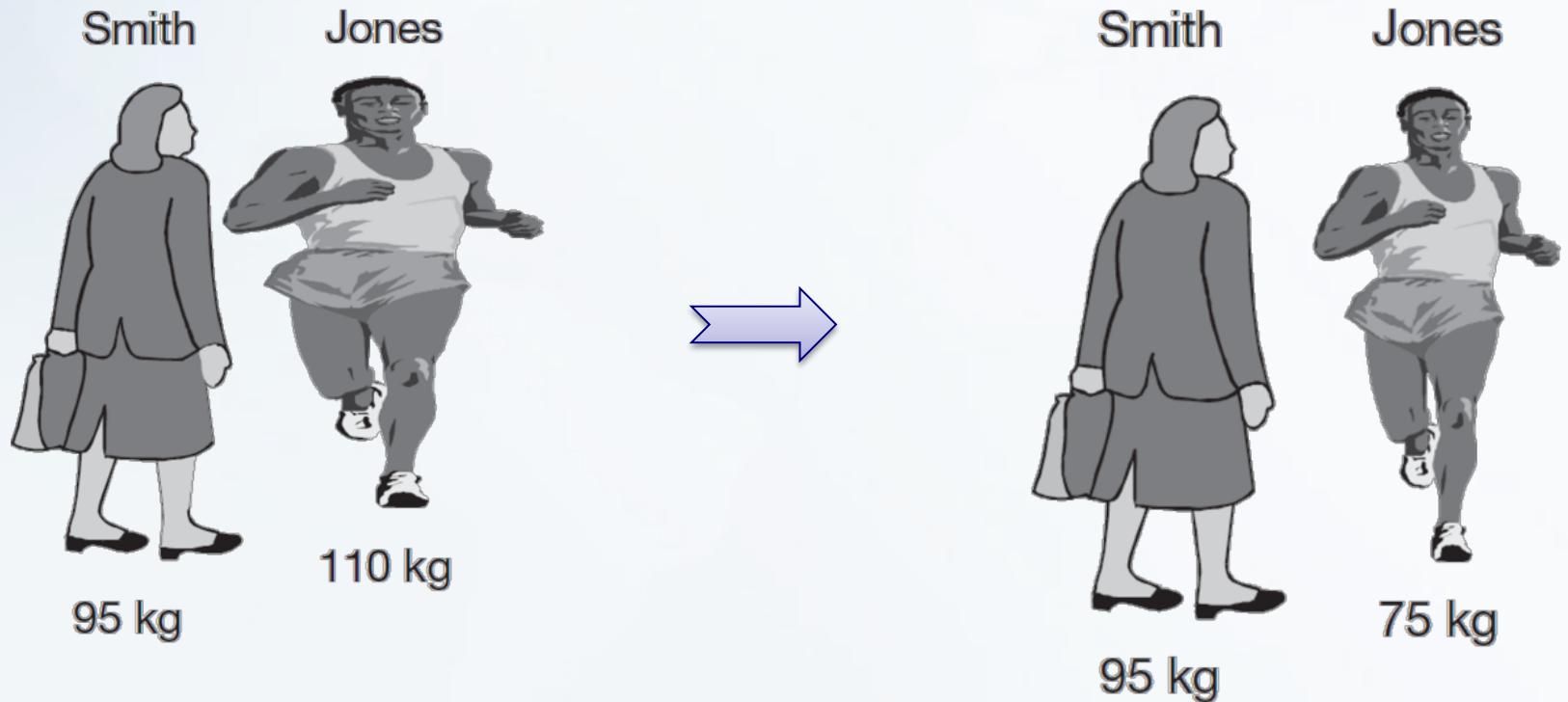
In Data Modeling:
- One of more attributes (Name, Name+Birthday, …) that we can use to **uniquely** refer to an entity, …within a contest!
- If we cannot easily find/use these attributes, we give an ID (ID number, Book number, URI, IRI ….)

In reality
- A property that can be used to **uniquely** refer to an entity… in any context, if this property is changed then the entity becomes another (what identifies a person? Book?).

Notice that an attribute/property should be mandatory/essential to be used for identification. (we will talk about this later).

What is Uniqueness?
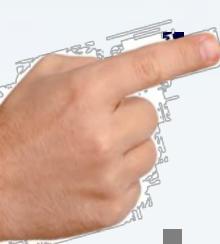
For each state taken individually, each person has at most one weight.

➜ How can we record such information without redundancy?

Uniqueness and Identity Rules in ORM
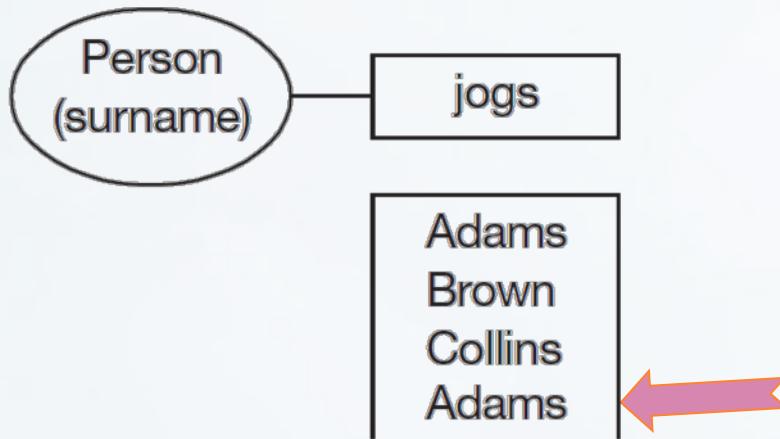
- Part 1: What is Uniqueness/Identity?

- **Part 2: Internal Uniqueness (within a Fact Type)**

- Part 3: External Uniqueness (across Fact Types)

- Part 4: Key Length and Reference Schema
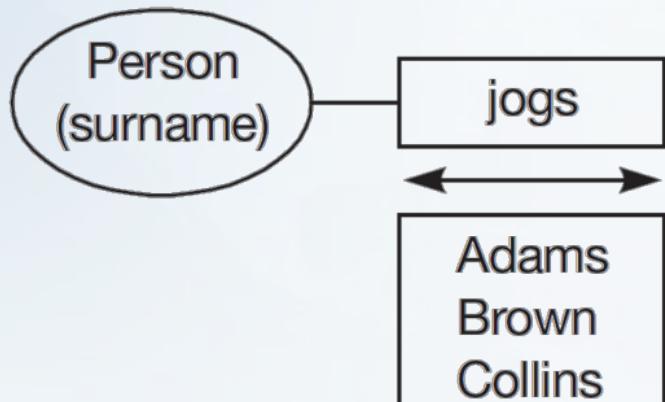
Uniqueness on Unary Fact Types
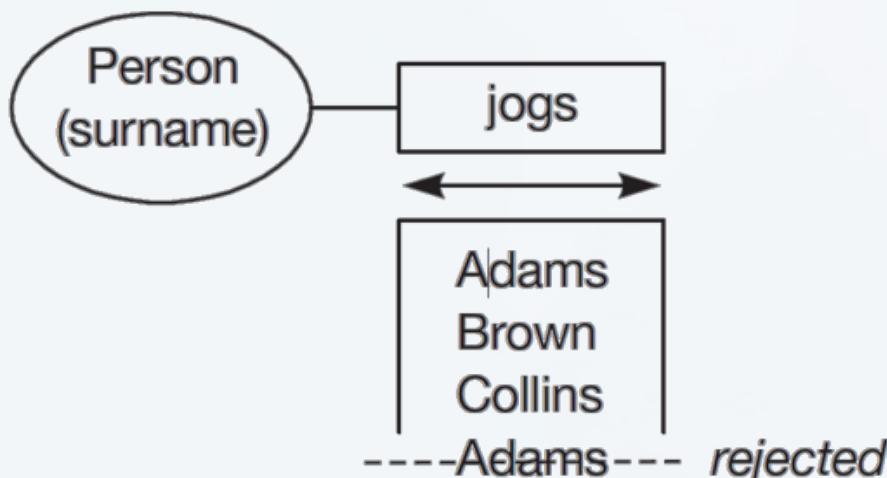
Is their any problem with this schema?

How can we prevent people adding such redundant information?

Uniqueness on Unary Fact Types

Person (surname) — jogs

⟷

Adams
Brown
Collins
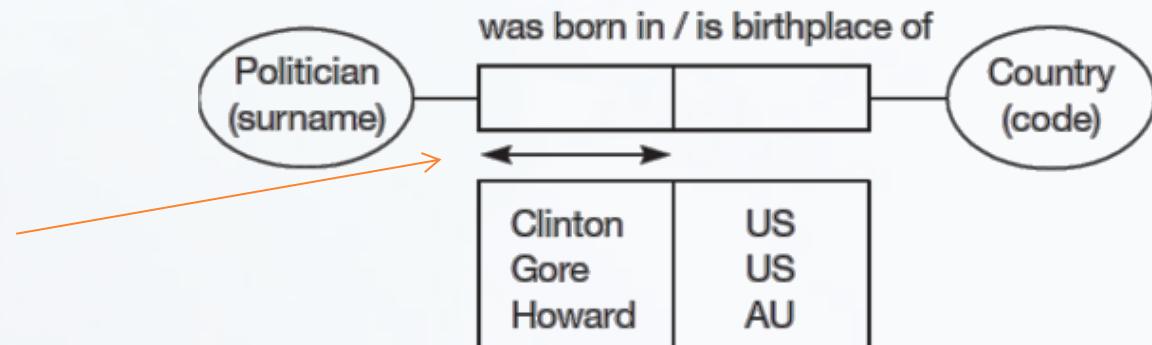
The uniqueness constraint ensures entities are unique (no duplicates)

Person (surname) — jogs

⟷

Adams
Brown
Collins
----Adams---- *rejected*
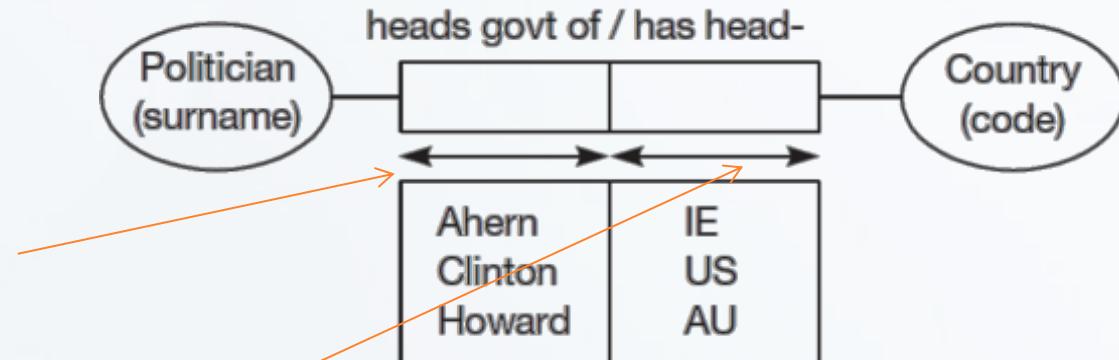
Uniqueness on Binary Fact Types

Each Politician was born in **at most one** Country

was born in / is birthplace of

Politician (surname)	Country (code)
Clinton	US
Gore	US
Howard	AU
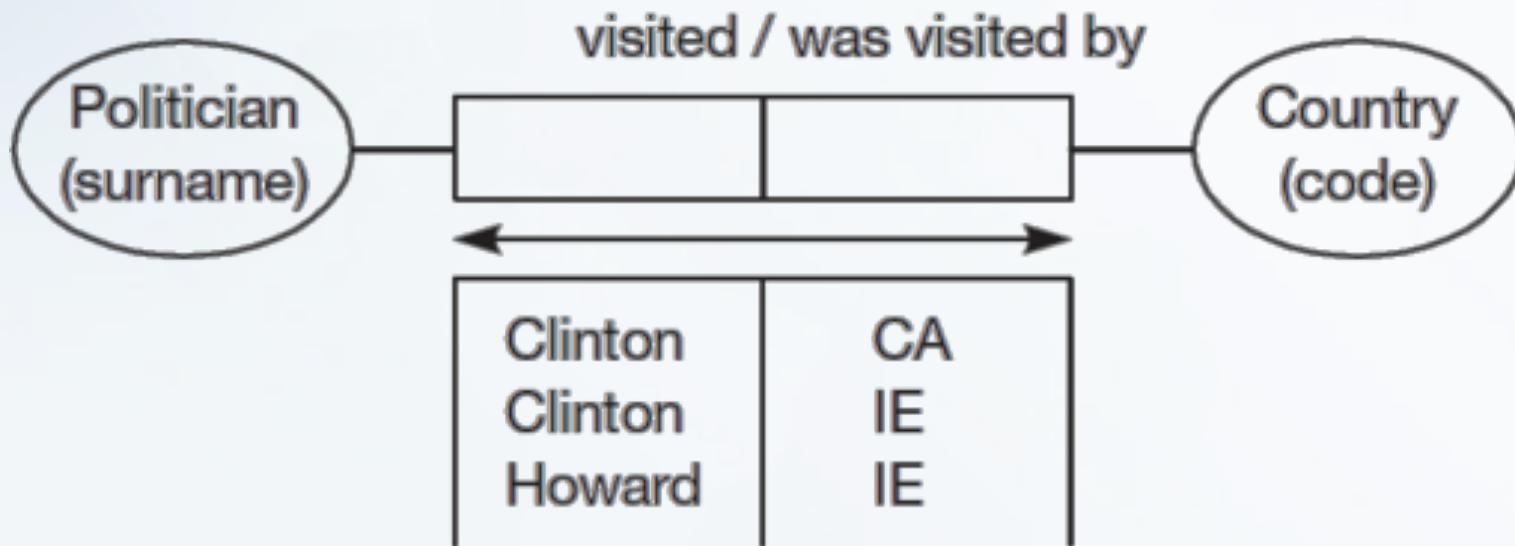
Each Politician heads government of **at most one** Country

Each Country has **at most one** head Politician

heads govt of / has head-

Politician (surname)	Country (code)
Ahern	IE
Clinton	US
Howard	AU

Uniqueness on Binary Fact Types

Means many to many

visited / was visited by

Politician (surname) — [|] — Country (code)

| Clinton | CA |
| Clinton | IE |
| Howard | IE |
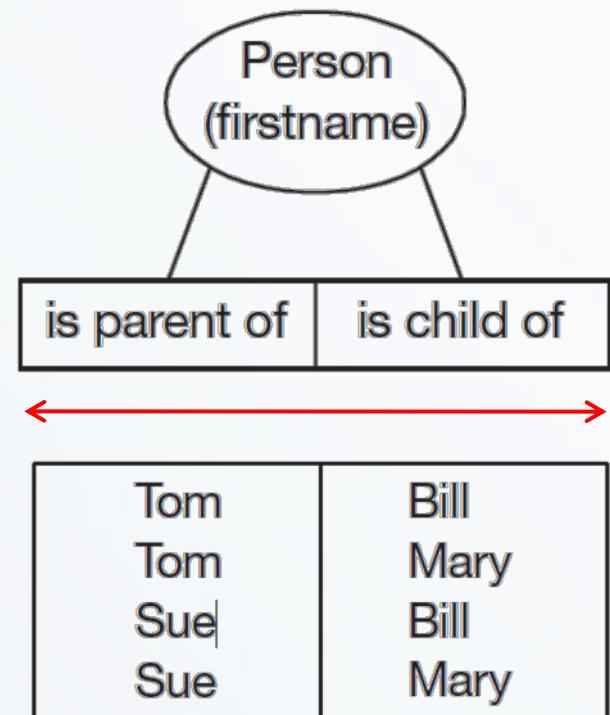
It is possible that the same Politician visited **more than one** Country **and that the same** Country was visited by **more than one** Politician

Who can give more examples?

Uniqueness on Binary Fact Types

What is unique here?

is parent of	is child of
Tom	Bill
Tom	Mary
Sue	Bill
Sue	Mary

Person (firstname)
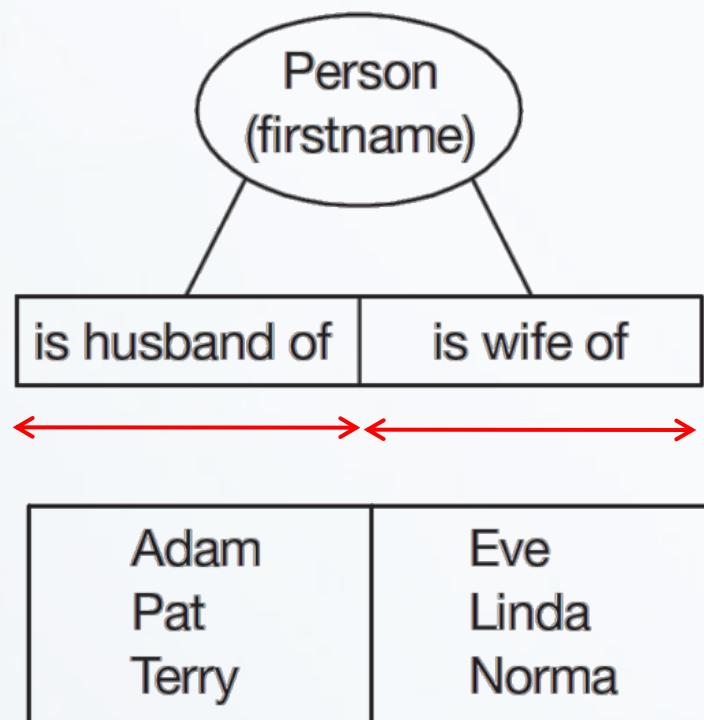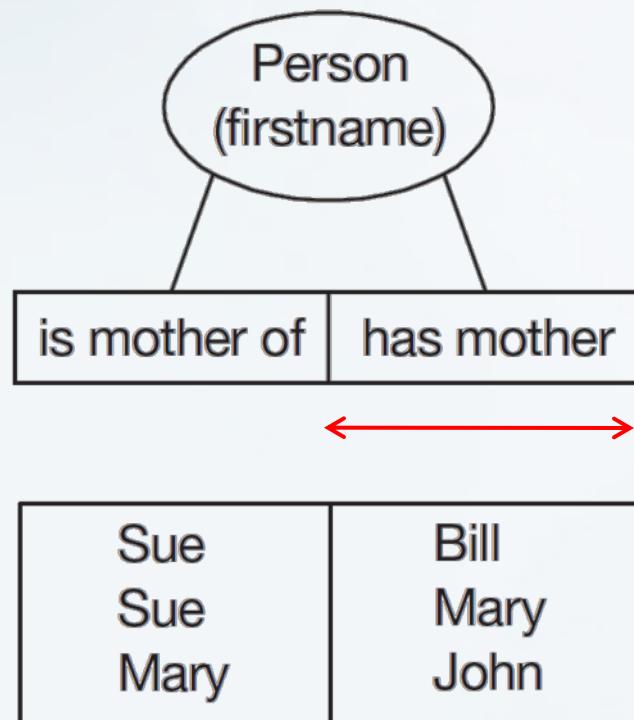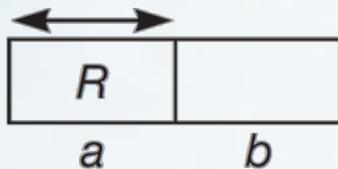
Uniqueness on Binary Fact Types

What is unique here?
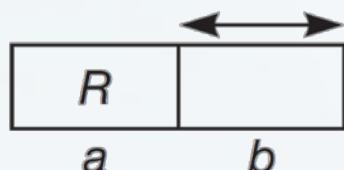
Uniqueness on Binary Fact Types

The four uniqueness constraint patterns for a binary.
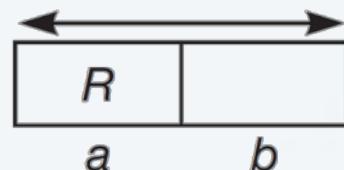
R | a b	No duplicates are allowed in **a**'s column Each **a R**'s at most one **b**
R | a b	No duplicates are allowed in **b**'s column Each **b** is **R**'d by at most one **a**
R | a b	Both the foregoing constraints apply
R | a b	No duplicate (**a**,**b**) rows are allowed Each **a** may **R** many **b** and vice versa
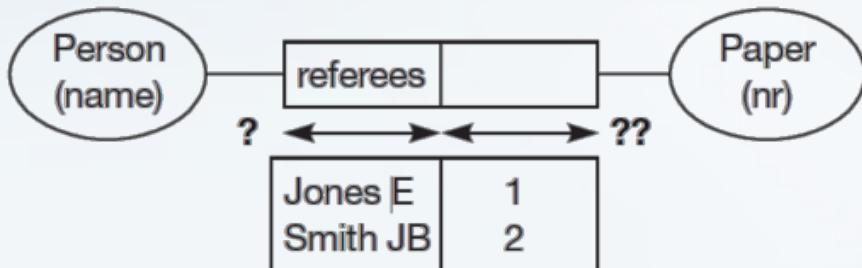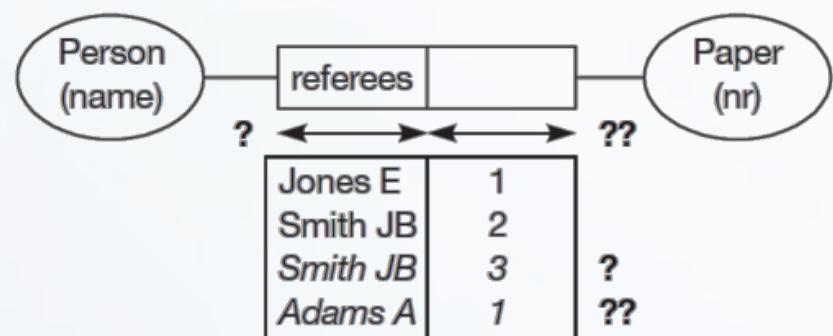
How to think about Uniqueness

| Referee	Paper Nr
Jones E	1
Smith JB	2
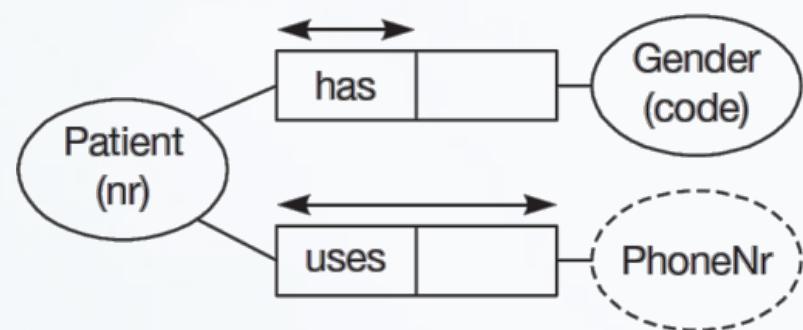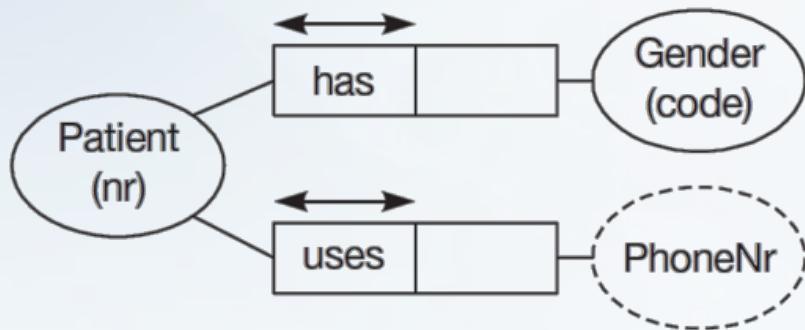
Is the population significant?

**Adding counterexamples to test the constraints**
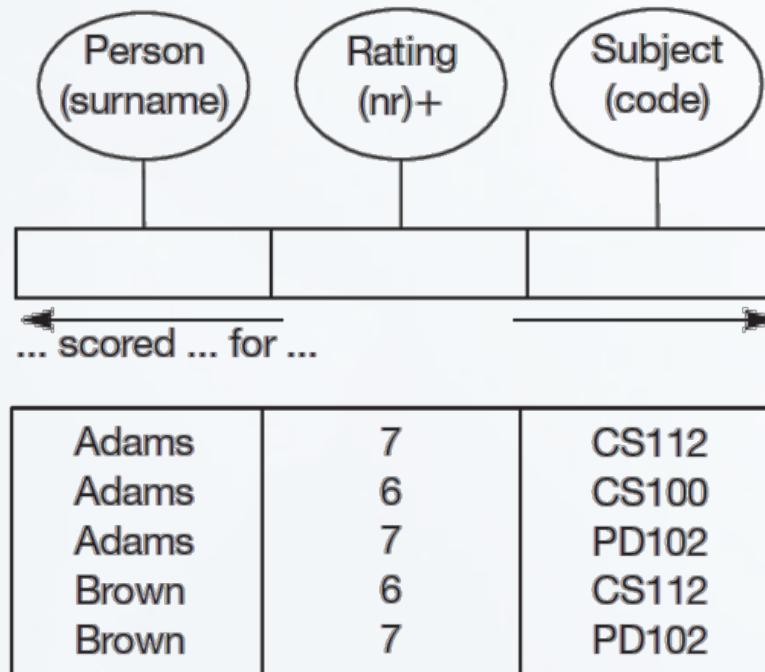
Uniqueness on Binary Fact Types


Which is more realistic?

Uniqueness on Ternary Fact Types
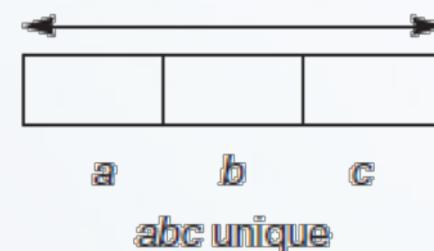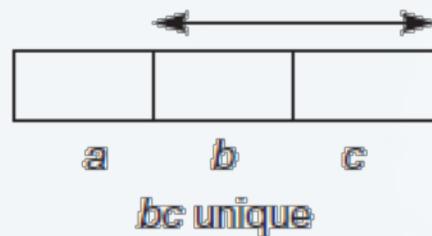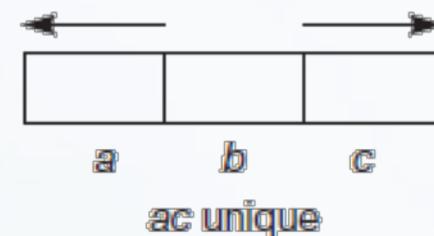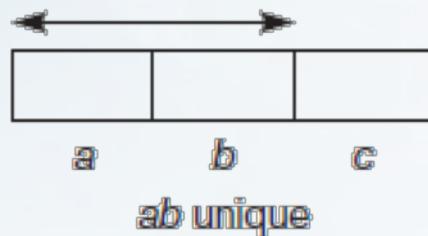
What are the uniqueness constraints?

Person (surname)	Rating (nr)+	Subject (code)
Adams	7	CS112
Adams	6	CS100
Adams	7	PD102
Brown	6	CS112
Brown	7	PD102

... scored ... for ...
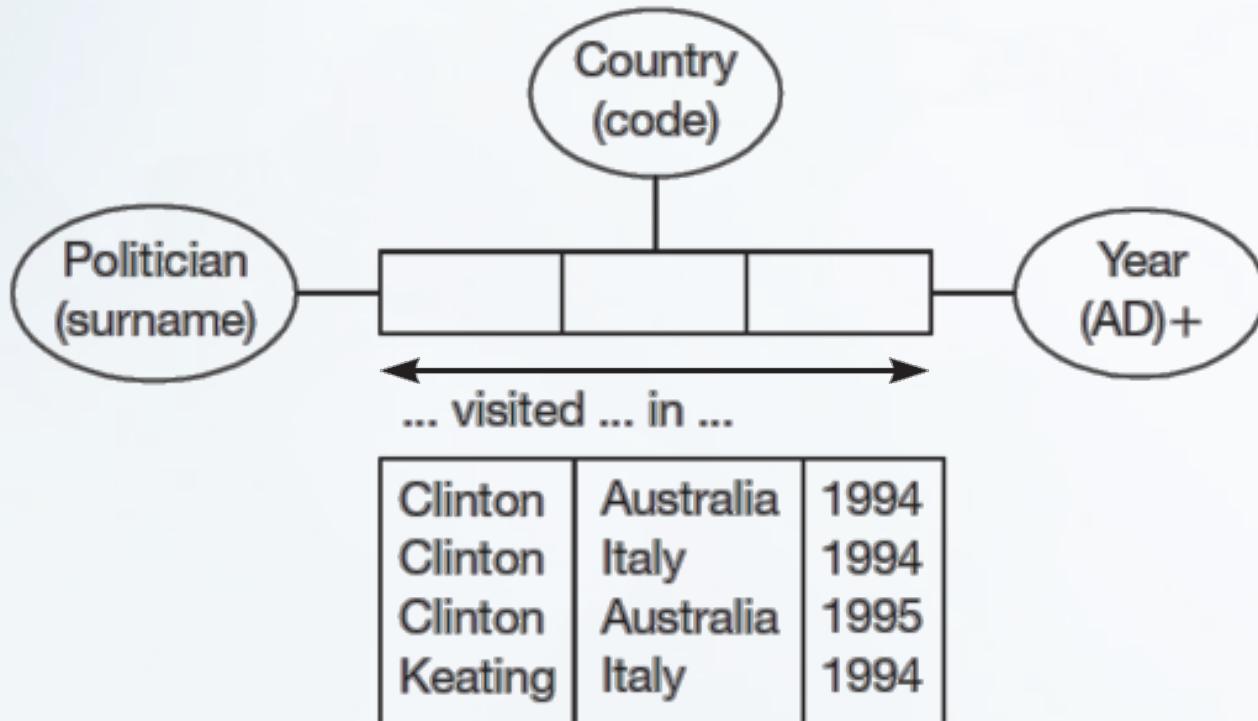
Each (Person, Subject) combination is unique.

Uniqueness on Ternary Fact Types

Allowed basic uniqueness constraints for a ternary.


ab unique

ac unique

bc unique

abc unique
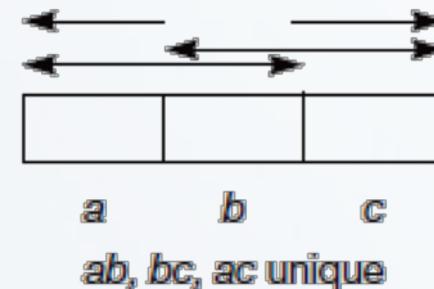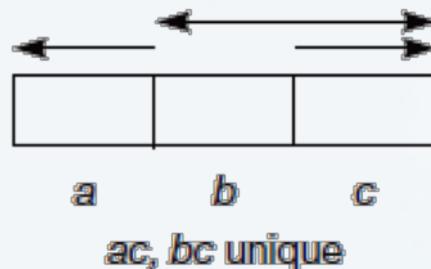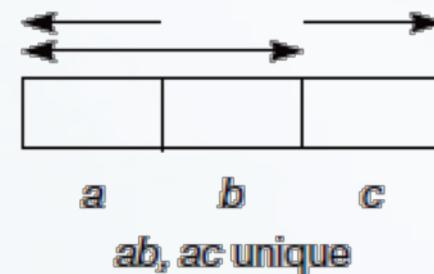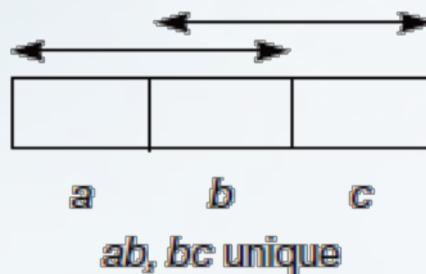
Uniqueness on Ternary Fact Types

What this uniqueness means?
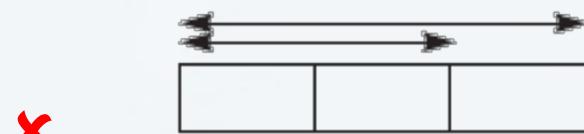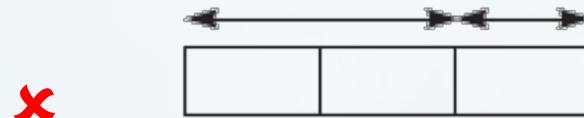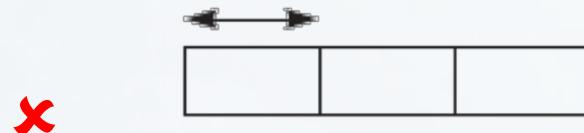
Uniqueness on Ternary Fact Types

Allowed uniqueness constraint combinations for a ternary.

ab, bc unique

ab, ac unique

ac, bc unique

ab, bc, ac unique
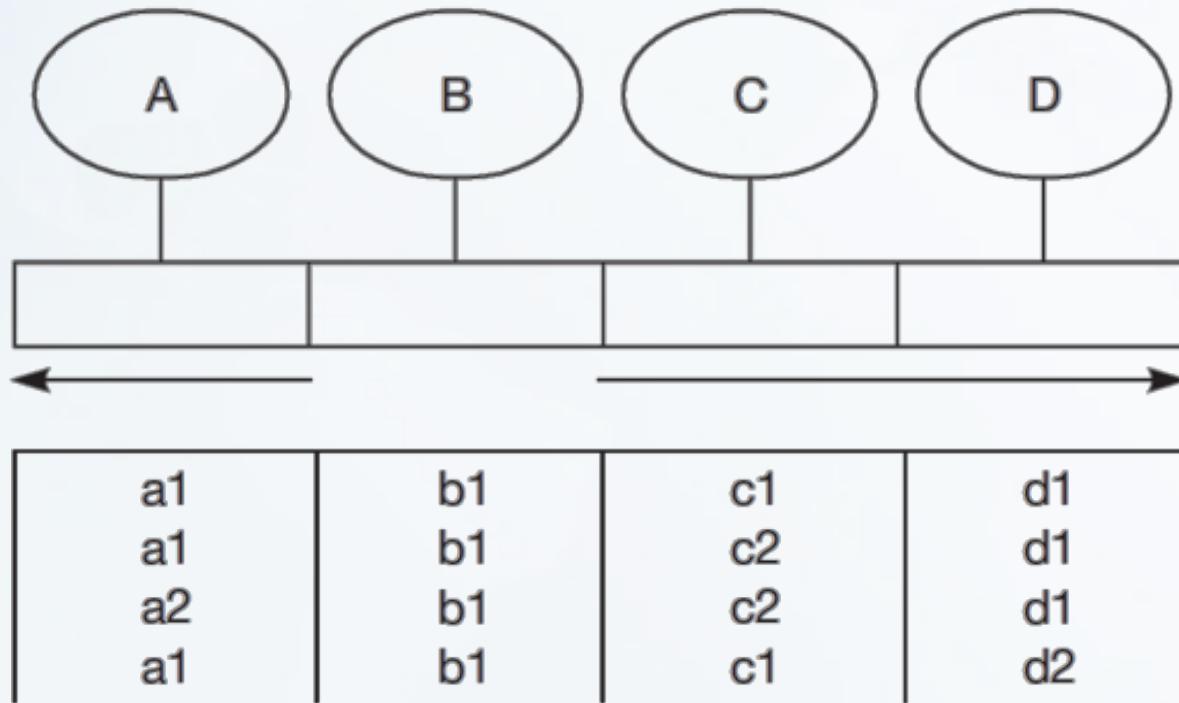
Uniqueness on Ternary Fact Types

Which of this constraint patterns is illegal? Why?
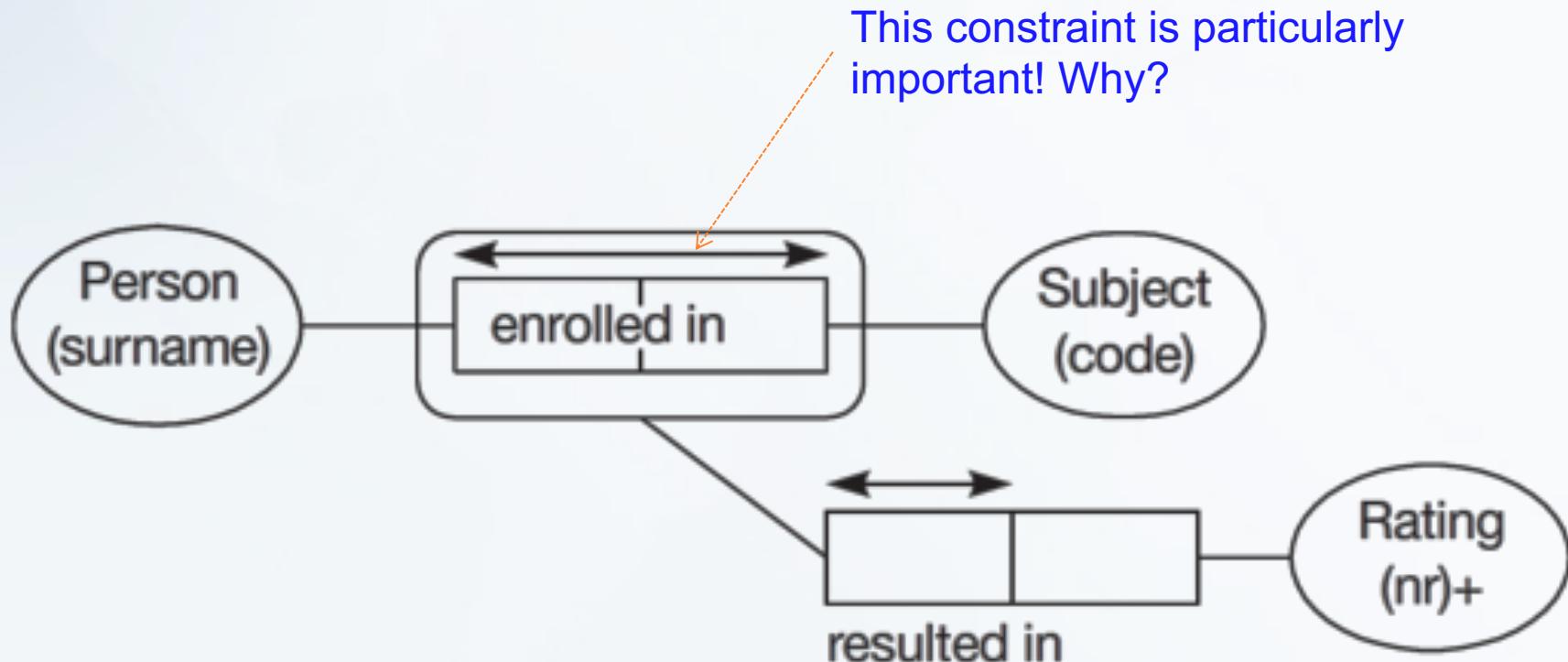
Example of Uniqueness on n-ary fact types

A	B	C	D
a1	b1	c1	d1
a1	b1	c2	d1
a2	b1	c2	d1
a1	b1	c1	d2

Each (a,c,d) combination occurs on at most one row.

Uniqueness with Nested Fact Types

This constraint is particularly important! Why?

➔ Explain what is unique

Uniqueness and Identity Rules in ORM
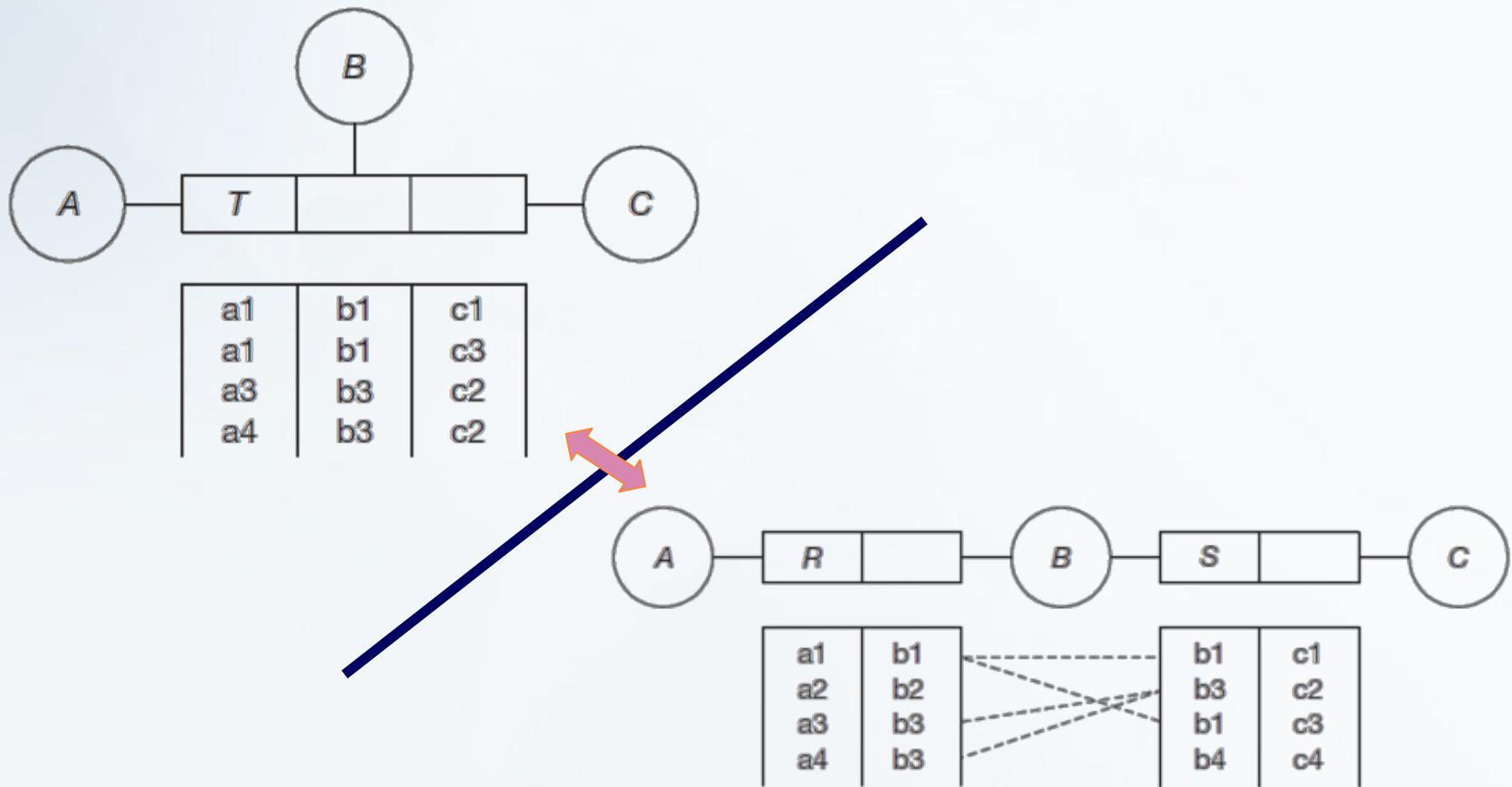
- Part 1: What is Uniqueness/Identity?

- Part 2: Internal Uniqueness (within a Fact Type)

- **Part 3: External Uniqueness (across Fact Types)**

- Part 4: Key Length and Reference Schema

What is the difference between these?
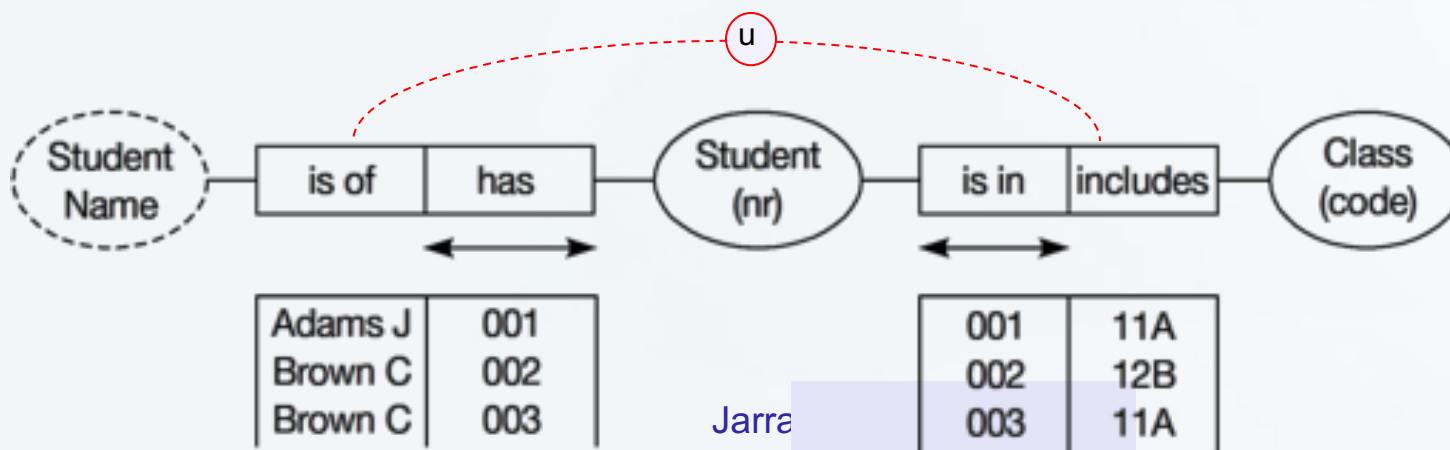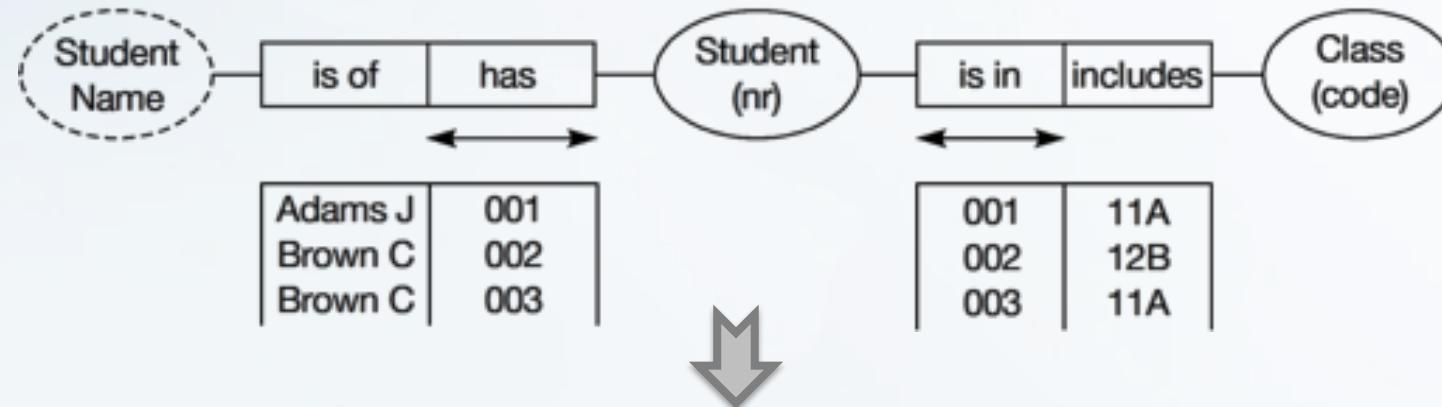
➔ Explain the Joins

➔ Do we need uniqueness?
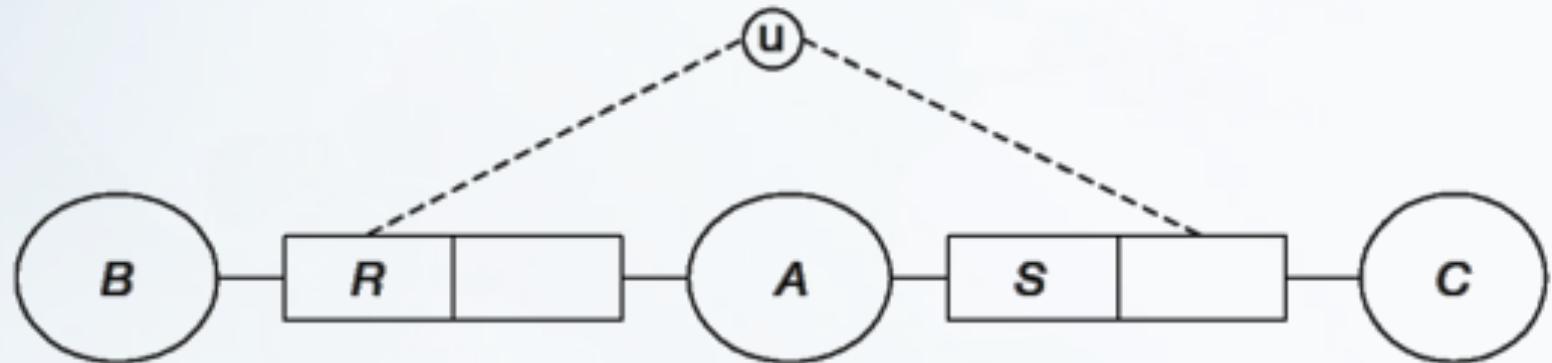
External Uniqueness constraints

StudentNr	Name	Class
001	Adams J	11A
002	Brown C	12B
003	Brown C	11A

What is missing?

External Uniqueness constraints
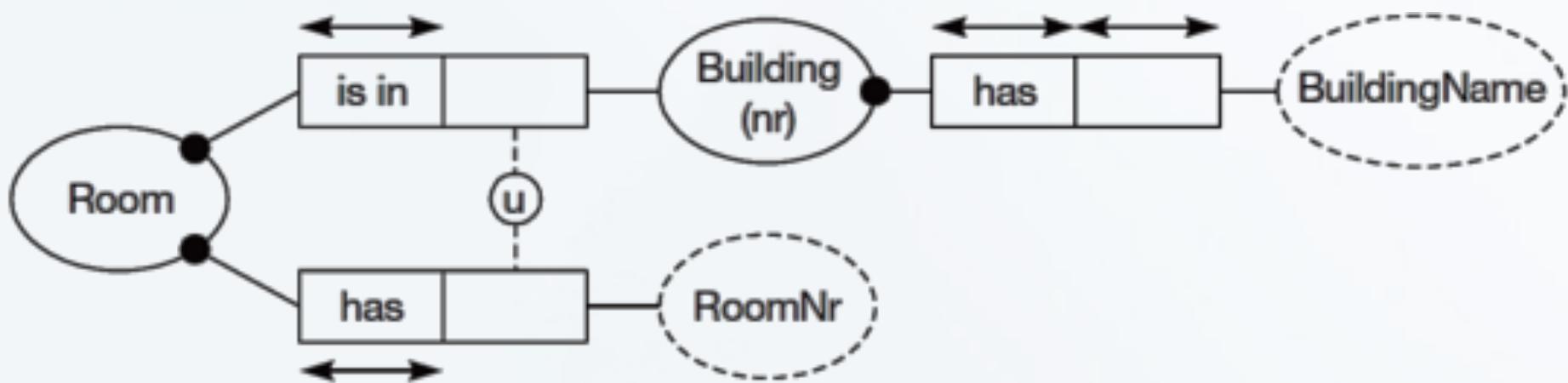
The meaning of the External Uniqueness

Each (**b**,**c**) combination is paired with at most one **a**

Each population **R** join **S** has **bc** unique
(where "join" denotes "conceptual inner join")
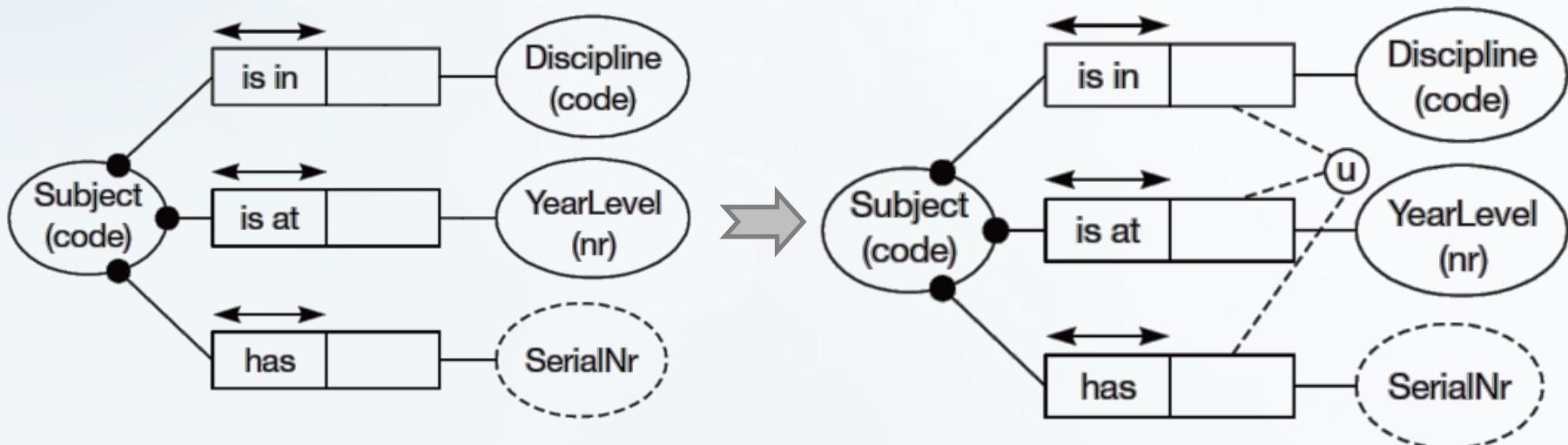
Example

BuildingNr	Building name	Nr rooms
...	...	...
67	Priestly	100
68	Chemistry	100
69	Computer Science	150
...	...	...
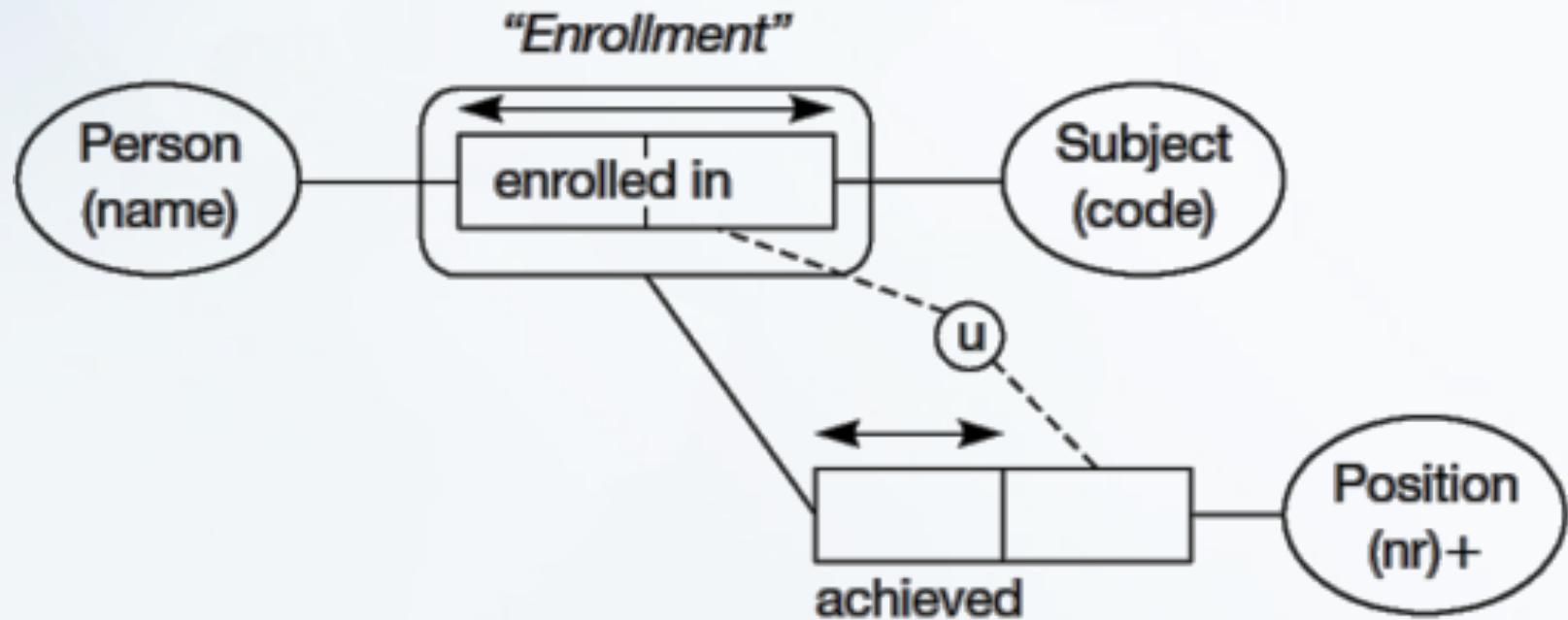
An Other Uniqueness

How to say that the combination of (Discipline, YearLevel, SerialNr) is unique for each subject?

The Subject code might be generated from this combination
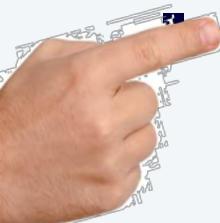
Example with nest fact types
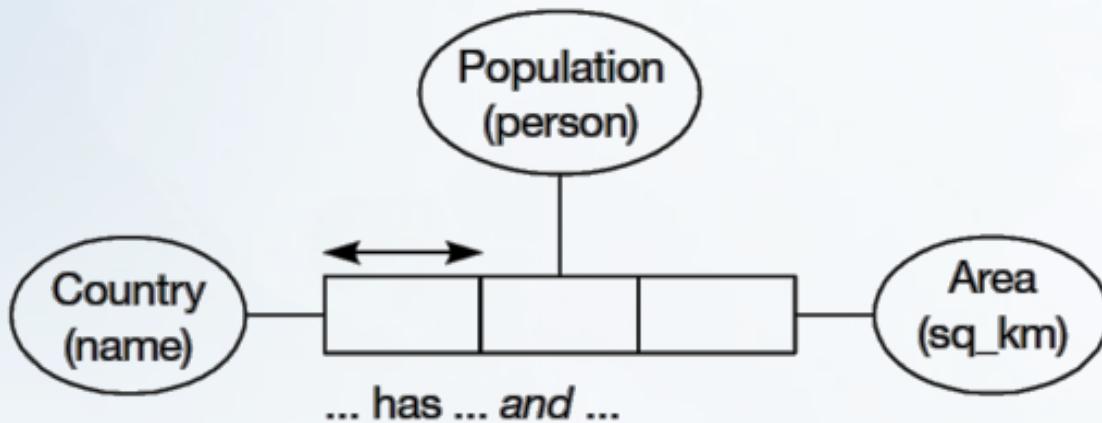
Uniqueness and Identity Rules in ORM

- Part 1: What is Uniqueness/Identity?

- Part 2: Internal Uniqueness (within a Fact Type)

- Part 3: External Uniqueness (across Fact Types)
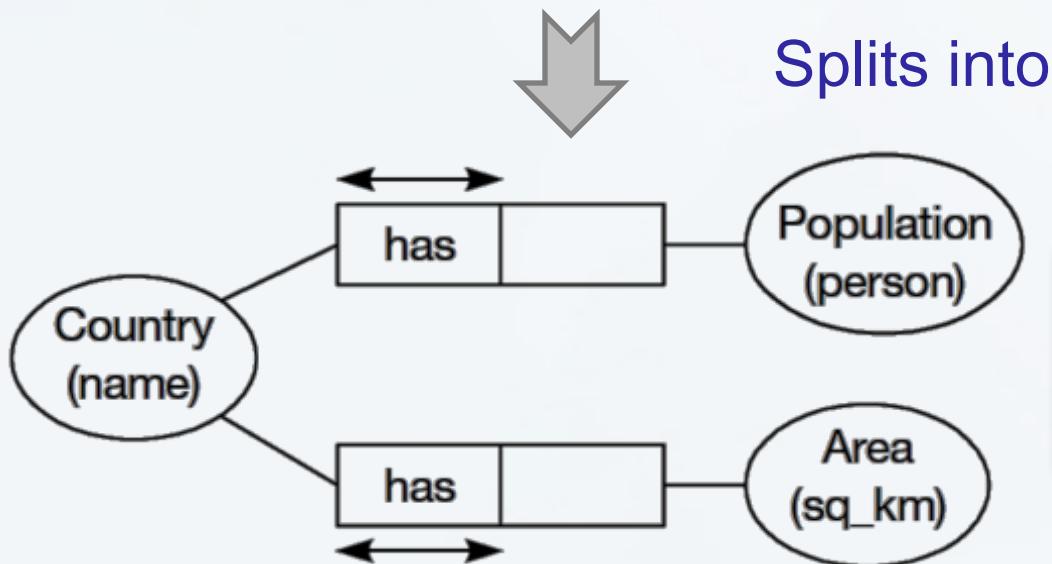
- **Part 4: Key Length and Reference Schema**

Key Length Check

... has ... *and* ...
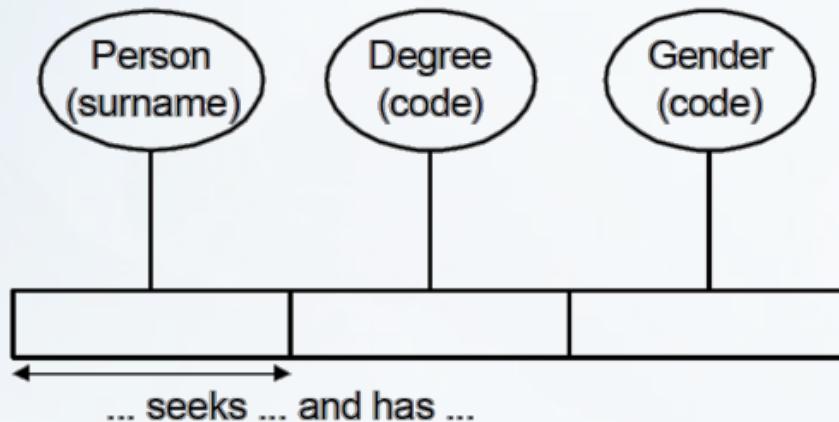
What is wrong?

Splits into

Each UC in an *elementary n-ary* relationship must span at least *n-1 roles*

Key Length Check
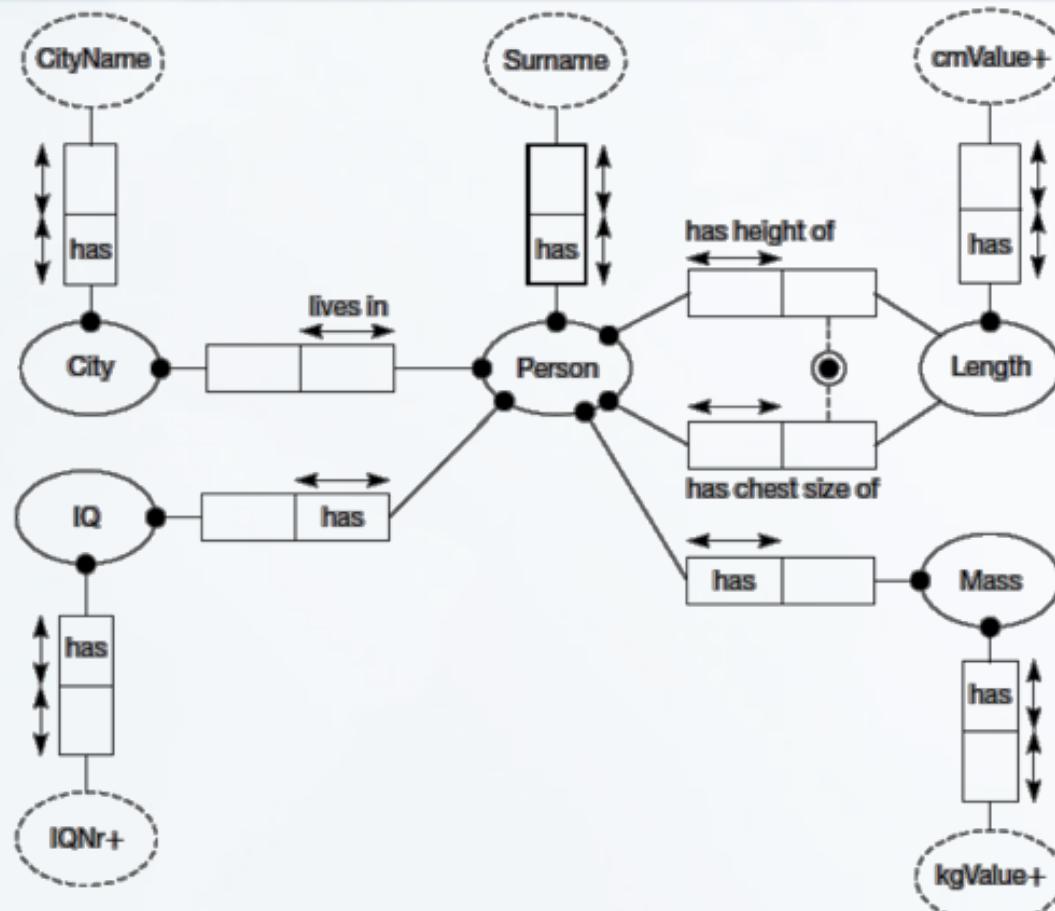
What is wrong?

Splits into

Reference Schemes

In data modeling, <u>each entity (i.e., Object Type) must have an identity</u>.

The identity is achieved if an entity has mandatory and unique role

➔This is not important (i.e. implicit) in ontology modeling.

Project (Student Registry)

Description:

The central management of students' profiles by the ministry of education is becoming an urgent need in the last years. Many students in Palestine move from one university to another, and they need to transfer their academic records. Also, the ministry of higher education needs to certify the diplomas and mark sheets of students. Moreover, there is a need to centrally manage/monitor students financial aids. Therefore, the ministry of higher education has decided to build a national student registry, such that, each semester every university has to send the academic record (i.e., mark-sheet) of every student to the ministry of education. The ministry will then update and integrate the academic records according to the data combined from all universities into the national student registry.

The ministry wants to specify a data model (in ORM) to be used as a reference data model of, including the business rules.

➔Develop a conceptual model (in both Arabic and English separately) for this information system, which must be suitable for mark sheets in all Palestinian universities (Tip: you may start with your own mark sheet).

➔ Deliver hard copy to my office, before February xx, 2018

References

[1] Terry Halpin, Tony Morgan: Information Modeling and Relational Databases, Second Edition. Second Edition. The Morgan Kaufmann Series in Data Management Systems. ISBN: 0123735688

[2] Mustafa Jarrar and Robert Meersman: Ontology Engineering -The DOGMA Approach. Book Chapter in "Advances in Web Semantics I". Chapter 3. Pages 7-34. LNCS 4891, Springer.ISBN:978-3540897835. (2008).

[3] Mustafa Jarrar, Anton Deik, Bilal Faraj: Ontology-Based Data And Process Governance Framework -The Case Of E-Government Interoperability In Palestine . In pre-proceedings of the IFIP International Symposium on Data-Driven Process Discovery and Analysis (SIMPDA'11). Pages(83-98). ISBN 978-88-903120-2-1. Campione, Italy. June 30, 2011.

[4] Mustafa Jarrar: Mapping ORM Into The SHOIN/OWL Description Logic- Towards A Methodological And Expressive Graphical Notation For Ontology Engineering . In OTM 2007 workshops: Proceedings of the International Workshop on Object-Role Modeling (ORM'07). Pages (729-741), LNCS 4805, Springer. ISBN: 9783540768890. Portogal. November, 2007

[5] Mustafa Jarrar: Towards Automated Reasoning On ORM Schemes. -Mapping ORM Into The DLR_idf Description Logic. In proceedings of the 26th International Conference on Conceptual Modeling (ER 2007). Pages (181-197). LNCS 4801, Springer. Auckland, New Zealand. ISBN 9783540755623. November 2007

[6] Mustafa Jarrar and Stijn Heymans: Unsatisfiability Reasoning In ORM Conceptual Schemes. In Current Trends in Database Technology - EDBT 2006: Proceeding of the IFIP-2.6 International Conference on Semantics of a Networked. Pages (517-534). LNCS 4254, Springer. Munich, Germany. ISBN: 3540467882. March 2006.

[7] Mustafa Jarrar and Stijn Heymans: Towards Pattern-Based Reasoning For Friendly Ontology Debugging . Journal of Artificial Intelligence Tools. Volume 17. No.4. World Scientific Publishing. August 2008.

[8] Mustafa Jarrar, Maria Keet, and Paolo Dongilli: Multilingual Verbalization Of ORM Conceptual Models And Axiomatized Ontologies. Technical report. STARLab, Vrije Universiteit Brussel, February 2006.

[9] Sergey Lukichev and Mustafa Jarrar: Graphical Notations For Rule Modeling . Book chapter in "Handbook of Research on Emerging Rule-Based Languages and Technologies". IGI Global. ISBN:1-60566-402-2. (2009)

[10] Mustafa Jarrar: Modularization And Automatic Composition Of Object-Role Modeling (ORM) Schemes .OTM 2005 Workshops: Proceedings of the Object-Role Modeling (ORM'05). Pages (613-625). LNCS 3762, Springer. ISBN: 3540297391. 2005.

[11] Mustafa Jarrar: Towards Methodological Principles For Ontology Engineering. PhD Thesis. Vrije Universiteit Brussel. (May 2005)

[12] Mustafa Jarrar, Jan Demey, and Robert Meersman: On Using Conceptual Data Modeling For Ontology Engineering . Journal on Data Semantics, Special issue on "Best papers from the ER/ODBASE/COOPIS 2002 Conferences". LNCS 2800. No 1. Springer. 2003.

[13] Jan Demey, Mustafa Jarrar, and Robert Meersman: A Markup Language For ORM Business Rules . Proceedings of the International Workshop on Rule Markup Languages for Business Rules on the Semantic Web (RuleML 2002). Pages(107-128). Volume 60. CEUR Workshop Proceedings. ISSN 1613-0073. June 2002

[14] Mustafa Jarrar: Towards Effectiveness And Transparency In E-Business Transactions, An Ontology For Customer Complaint Management . A book chapter in "Semantic Web Methodologies for E-Business Applications". chapter 7. IGI Global. (2008)

[15] Mustafa Jarrar: ORM Markup Language, Version 3 . Technical Report. STAR Lab, Vrije Universiteit Brussel, Belgium. January 2007