

# Data Integration: Metadata

Jan Chomicki

University at Buffalo

## Schematic discrepancies

The information in the **schema** of one database may correspond to the information in the **instance** of another database.

### Postulates

- 1 the same constant may be a relation name, a column name and an attribute value
- 2 schema elements should be first-class objects
- 3 queries may define more than one relation, each with varying number of columns.

### Correspondences

- attribute value  $\iff$  relation name
- attribute value  $\iff$  column name
- column name  $\iff$  relation name

A successor of SchemaSQL [LSS01].

## Features

- **metavariables**, ranging over relation and column names
- dynamically varying relation schemas (INTO, ON)

## Database schemas

DA: Exams(Sid,Exam,Grade)

DB: Theory(Sid,Grade), AI(Sid,Grade), Systems(Sid,Grade)

DC: Students(Sid,Theory,AI,Systems)

Attribute values  $\iff$  relation names

## DA2DB

```
select E.Sid as "Sid", E.Grade as "Grade"  
into E. Exam  
from DA.Exams as E
```

## DB2DA

```
select T.Sid as "Sid", R as "Exam", T.Grade as "Grade"  
into "Exams"  
from DB :R as T
```

## Attribute values $\iff$ column names

### DA2DC

```
select E.Sid as "Sid", E.Grade on E.Exam
into "Students"
from DA.Exams as E
```

### DC2DA

```
select C.Sid as "Sid", A as "Exam", C.A as "Grade"
into "Exams"
from DC: R as C, R:A
where R="Students" and A <> "Sid"
```

## FISQL queries

### Semantics

- a generalization of SQL semantics
- metavariables range over relation and column names
- output: special treatment for dynamic schemas

### FIRA

- extension of relational algebra
- operators map federated databases to federated databases
- new operators: partition, transpose,....

FISQL can be translated to FIRA and vice versa.



L.V.S. Lakshmanan, F. Sadri, and I.N. Subramanian.  
SchemaSQL – A Language for Interoperability in Relational  
Multi-Database Systems.  
*ACM Transactions on Database Systems*, 26(4):476–519, 2001.



C.M. Wyss and E. L. Robertson.  
Relational Languages for Metadata Integration.  
*ACM Transactions on Database Systems*, 30(2):624–660, 2005.