

# Schema Mappings and Data Examples: Deriving Syntax from Semantics

Phokion G. Kolaitis<sup>1</sup>

<sup>1</sup> University of California Santa Cruz & IBM Research - Almaden  
California, USA  
kolaitis@cs.ucsc.edu

---

## Abstract

Schema mappings are high-level specifications that describe the relationship between two database schemas. Schema mappings are considered to be the essential building blocks in such critical data interoperability tasks as data exchange and data integration. For this reason, they have been the focus of extensive research investigations over the past several years. Since in real-life applications schema mappings can be quite complex, it is important to develop methods and tools for illustrating, explaining, and deriving schema mappings. A promising approach to this effect is to use “good” data examples that illustrate the schema mapping at hand.

In this talk, we present an overview of recent work on characterizing and deriving schema mappings via a finite set of data examples. We show that every LAV schema mapping (i.e., a schema mapping specified by a finite set of local-as-view tuple-generating dependencies) is uniquely characterized by a finite set of universal data examples with respect to the class of all LAV schema mappings. We also show that this type of result does not hold for arbitrary GAV schema mappings (i.e., schema mappings specified by a finite set of global-as-view tuple-generating dependencies). After this, we give a necessary and sufficient algorithmic condition for a GAV schema mapping to be uniquely characterizable by a finite set of universal examples with respect to the class of all GAV schema mappings. Along the way, we establish tight connections between unique characterizability of schema mappings and homomorphism dualities.

This is joint work with Bogdan Alexe (IBM Research - Almaden), Balder ten Cate (UC Santa Cruz), and Wang-Chiew Tan (UC Santa Cruz and IBM Research - Almaden) based on [1, 2, 3].

**1998 ACM Subject Classification** D.2.12 Interoperability: Data mapping; H.2.5 Heterogeneous Databases: Data translation

**Keywords and phrases** Schema mappings, database constraints, data exchange, data integration, universal solutions, homomorphism dualities

**Digital Object Identifier** 10.4230/LIPIcs.FSTTCS.2011.25

---

## References

- 1 Bogdan Alexe, Phokion G. Kolaitis, and Wang-Chiew Tan. Characterizing schema mappings via data examples. In *ACM Symposium on Principles of Database Systems (PODS)*, pages 261–272, 2010.
- 2 Bogdan Alexe, Balder ten Cate, Phokion G. Kolaitis, and Wang-Chiew Tan. Designing and refining schema mappings via data examples. In *SIGMOD Conference*, pages 133–144, 2011.
- 3 Balder ten Cate, Phokion G. Kolaitis, and Wang-Chiew Tan. Database constraints and homomorphism dualities. In *International Conference on Principles and Practice of Constraint Programming (CP)*, pages 475–490, 2010.



© Phokion G. Kolaitis;

licensed under Creative Commons License NC-ND

31<sup>st</sup> Int'l Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2011).

Editors: Supratik Chakraborty, Amit Kumar; pp. 25–25

Leibniz International Proceedings in Informatics

**LIPICS** Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany