

SQL for SRL: Model Structure Learning Inside a Database System

or, Structure Learning Made Easy

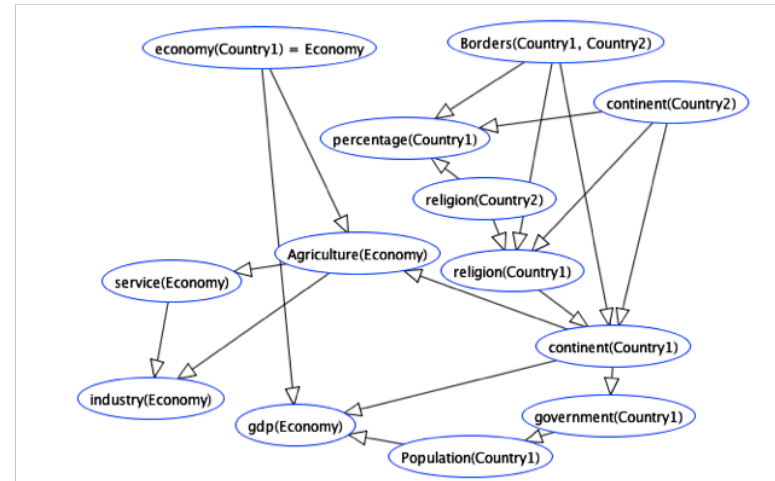
Oliver Schulte and Zhensong Qian
Simon Fraser University, Vancouver, Canada.

The logo for Simon Fraser University (SFU), consisting of the letters "SFU" in white, bold, sans-serif font, centered within a solid red rectangular background.

SFU

SQL for Programming Model Structure Learning

- ▶ *Programming* graphical model search for relational data is hard.
 - Multi-relational data is not self-describing.
 - Need to query metadata. e.g. DB schema/mode declarations.
 - *Structured* Models (Graphs) with *structured* components (terms, predicates, first-order variables, constants).
 - Computing event *counts* across multiple tables is expensive and error-prone.
 - Large Parameter space, > 1M sometimes.



The Solution: SQL Scripts All the Way

- ▶ Store relational **model** inside the database.
(As well as relational data.)
Tuffy, Felix, BayesStore, Wang et al. 1995.
- ▶ Use SQL to query metadata from DB catalog.
- ▶ Native SQL support for complex counts (count(*)).
- ▶ SQL for creating, transforming, storing sets of models.
- ▶ SQL for computing and storing parameter values.
>1M parameters no problem.
- ▶ SQL is standardized: system is portable, works out of the box.

Related Work

- ▶ Tuffy, BayesStore: complementary
 - push model inside the database as well.
 - leverage database techniques for *inference*/parameter learning, not structure learning.
- ▶ Madlib, MLBase, Bismarck, MauveDB, Unipivot...
 - leverage database techniques for *single-table* learning.

