

CIS RESEARCH SEMINAR

Date: Thursday, October 11

Time: 3PM

Location: Levine 307

Speaker: Val Tannen

Title: "Data Provenance and Query Languages"

Abstract:

Data provenance is the history of "ownership" of various pieces of information involved in a scientific study or a business data warehouse. Such data is sometimes mapped from data sources through views that can be described using query languages. In this case, scientists or business intelligencers are interested in also tracking through the views the provenance of the data.

Using a novel framework based on a bit of ancient algebra, we discuss provenance tracking of relational data through relational algebra and datalog views.

It turns out that this also generalizes nicely calculations done for incomplete and probabilistic databases, as well as data multiplicity.

-----