Database Metatheory: Asking the Big Queries

Christos H. Papadimitriou

Discussion: Ashique
Presentation: Jerry

Introduction

- What is theory? In general:
  - Abstraction: suppression of low-level details
  - Goal: see fundamental truths obscured by details
- In CS, theory is generally mathematical:
  - Developing models
  - Using model
  - Analyzing models

How does one do theory?

- Develop Mathematical Models
  - E.g., Turing machines
- Propose Complexity-Reducing Solutions
  - E.g., algs. for answering queries using views
- Analyze
  - E.g., transaction processing
- Explore!
  - What are real semantics of NULL?

The Joys and Pains of Exploration

- Joys:
  - Historically useful
  - In reasonable amounts, ensures good health
  - Theories are pretty: people will do it anyway
- Pains:
  - Must not consistently ignore practice
  - Requires careful exposition of relevance and applicability
  - Too much can lead to crises

What is "Good Theory"

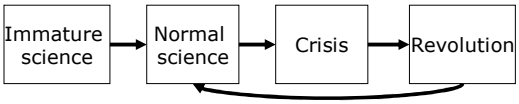
- *All* ideas improve knowledge
- But whether it's "good" theory largely depends on propaganda
  - Needs to influence beyond itself
  - Has to at least *be able to* influence practice

The ultimate influence: launching a victorious scientific revolution

Discussion 1

- In section 4 we saw experimental validation helps a theory to be successful. How important is theoretical validation for practices to succeed?
- There are some strange situations where a practice badly needs a theoretical validation; I can give an example like homeopathy, an alternative medicine that's been round the corner for 200 years but without theoretical validation it hasn't got universal acceptance. So I would like to exchange thoughts and views with you if you know similar incidents and your take on how theory is important for getting practices acceptability.

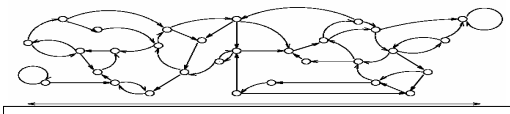
On Paradigms and Revolution (Thomas Kuhn's Model)



- "Normal" science has a predominant *paradigm*
  - Scientists pressured to defend paradigm and show it works
- Eventually, a crisis may (or may not) cause a revolution
  - E.g., object-oriented model & relational model

What's theory's role?

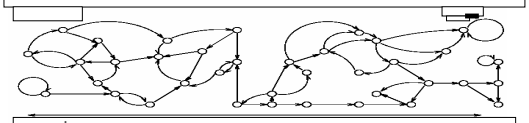
Theory's role in revolution: normal



practice theory

- Lots of connections
- Most theory within a few hops of practice, and vice-versa
- exploratory activities help fill uncharted regions and bring them to practice

Theory's role in revolution: crisis



practice theory

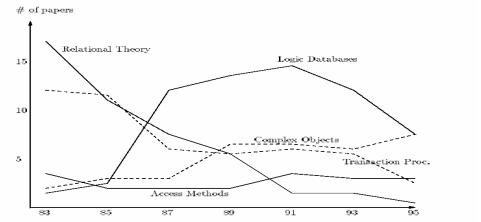
- Long paths from theory to practice
- Some nodes have no or little routes to practice
- In the short term, not good!
- In the long term, can help create new paradigm and new practice

Discussion 2

With respect to Kuhn's model, in what state is each of the following areas of database research? Normal? Crisis? Revolution? Justify your choice.

- relational roots
- query optimization
- query execution
- transaction processing
- extensible databases
- distributed databases
- views
- adaptive execution
- object oriented DBs
- XML
- data mining
- streaming data
- DB administration

What about database theory? (as seen by PODS papers)



- In the beginning (1982), there was relational theory and transaction processing
- Then datalog, objects, XML (not shown)

How did database theory do?

- Big Win:
  - Relational model & normal forms
- Big Loss:
  - Datalog & recursive queries (a bit better now)
- Draws:
  - Object-oriented models?
  - Only simplest concurrency control used

Dangerous Applicability Claims

- Recursive applicability
  - The last n papers said it was applicable
- Remote applicability
  - People in other fields find it applicable
- Applicability by association
  - If X is relevant to Y, then anything involving X must be applicable
- Applicability by pun
  - Since X and Y are important , X + Y must be so

Discussion 3

"a theory is a good theory if it satisfies two requirements: It must accurately describe a large class of observations on the basis of a model that contains only a few arbitrary elements, and it must make definite predictions about the results of future observations."

-- Stephen Hawking in 'A Brief History of Time'

This comment contradicts with what Feyerabend said about good theory. What makes good theory? Scientific merit? Applicability? Propaganda?