Database Metatheory: Asking the Big Queries

:Presenters
Shervin Mohammadi
Valerie Ishida

Discussion Leader:
Shriram Rajagopalan

What is theory?

- In Context of applied Science:
  - Theory is the use of significant abstraction in scientific research (in the board sense)
  - Abstraction: suppression of low-level details of the object or artifact being studied or designed.
- In Computer Science:
  - Theory is the use of sophisticated mathematical techniques for:
    - **Developing** Mathematical Models
    - **Using** Mathematical Models
    - **Analyzing** Mathematical Models

Theory is needed to cure **COMPLEXITY**

How theoreticians attack the COMPLEXITY

3)Develop mathematical models of the artifact

- Turing machine, relational model

4)Propose complexity-reducing solutions

- Algorithms and representational schemes

(e.g. Algorithms for answering queries using views)

5)Analyze the mathematical model

- Predict outcomes of experiments
- Calibrate models

6)Explore

- What is “important” and “relevant”

Exploration

Benefits:

Explorations

- Have been Historically beneficial
- Promote the field's health and connectivity in reasonable amounts
- Are natural and attractive activities

Dangers:

Explorations

- Can disoriented the field and lead to crisis
- Will not thrive if it consistently ignores practice
- Require precise discipline in its exposition of relevance and applicability

Discussion-1

"The devil is in the details". Theoretical research neglects the details and focuses on the abstract model of the real world scenario.

This is the reason why practioners neglect the outputs of theoretical research. --Agree ?

But such theoretical research helped in advancement of OS (Multitasking, scheduling), Security (cryptography), Search (information retrieval).
Defend.

(or)

Defend: Details dilute the problem being studied. They change with their environment. Only abstract research has caused fundamental changes in science (creation of computers for e.g.)

What is "Good Theory"?

Paul Feyerabend:

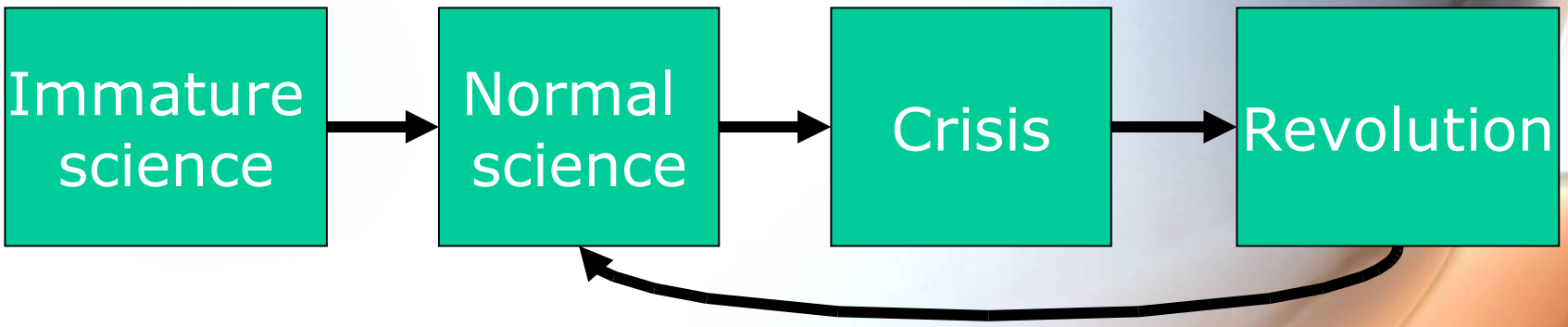
"[...] There is no idea that is not capable of improving our knowledge. [...] The only principle that doesn't inhibit progress is 'anything goes'."

- All ideas improve knowledge

- Theoretical Research's success mainly depends on its **Propagandistic Value:**
  - Contaminate its environment
  - Invade and Affect other ideas, fields
  - Influence practical milieu

- *"The ultimate success of scientific idea is the launching a victorious **scientific revolution**"*

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 causes a revolution
  - E.g., relational model

What's theory's role?

Discussion

"Scientific Progress in a direction. Crisis when critical mass is reached by neglected problems of the current paradigm. Finally THEORY CAUSES a Paradigm Shift - Revolution."

Note: Paradigm shift here means a SIGNIFICANT CHANGE, not minor.

Do you think:

This pattern occurs in Programming Languages, Operating Systems and other areas in computer science too?

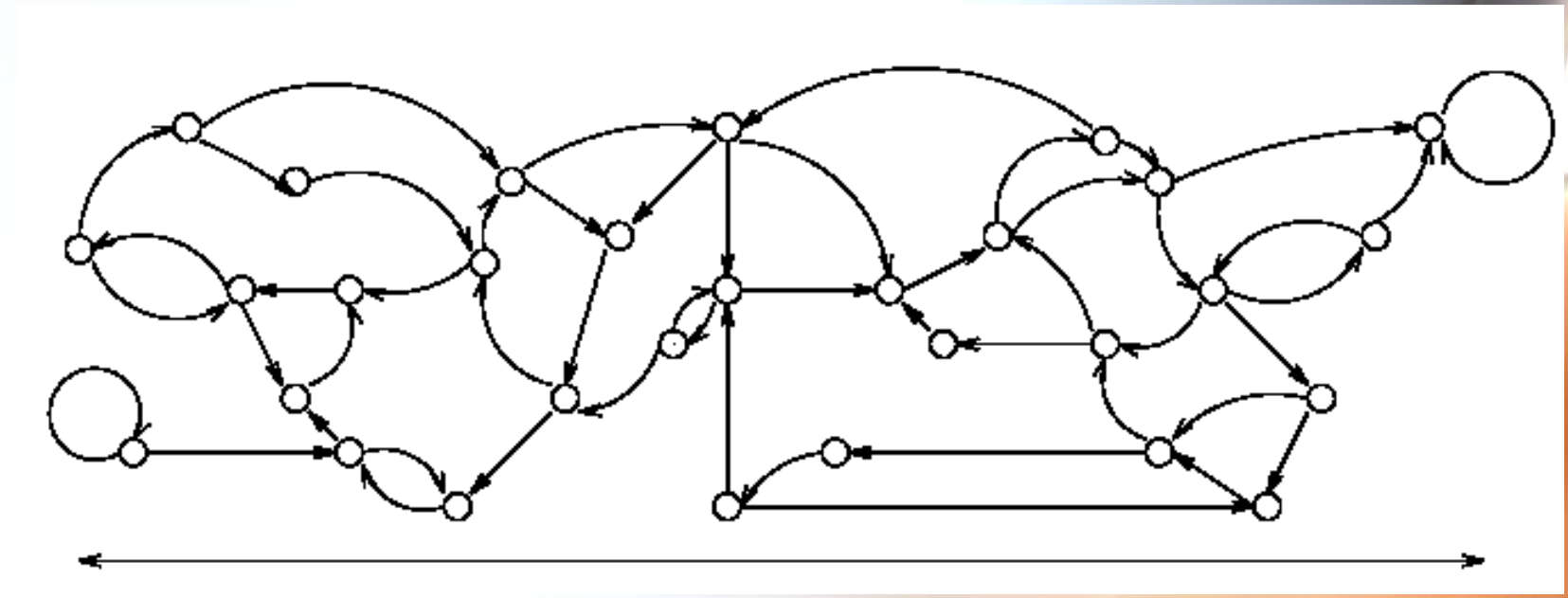
or

Paradigm shifts did occur but it was not theory that caused these paradigm shifts?

or

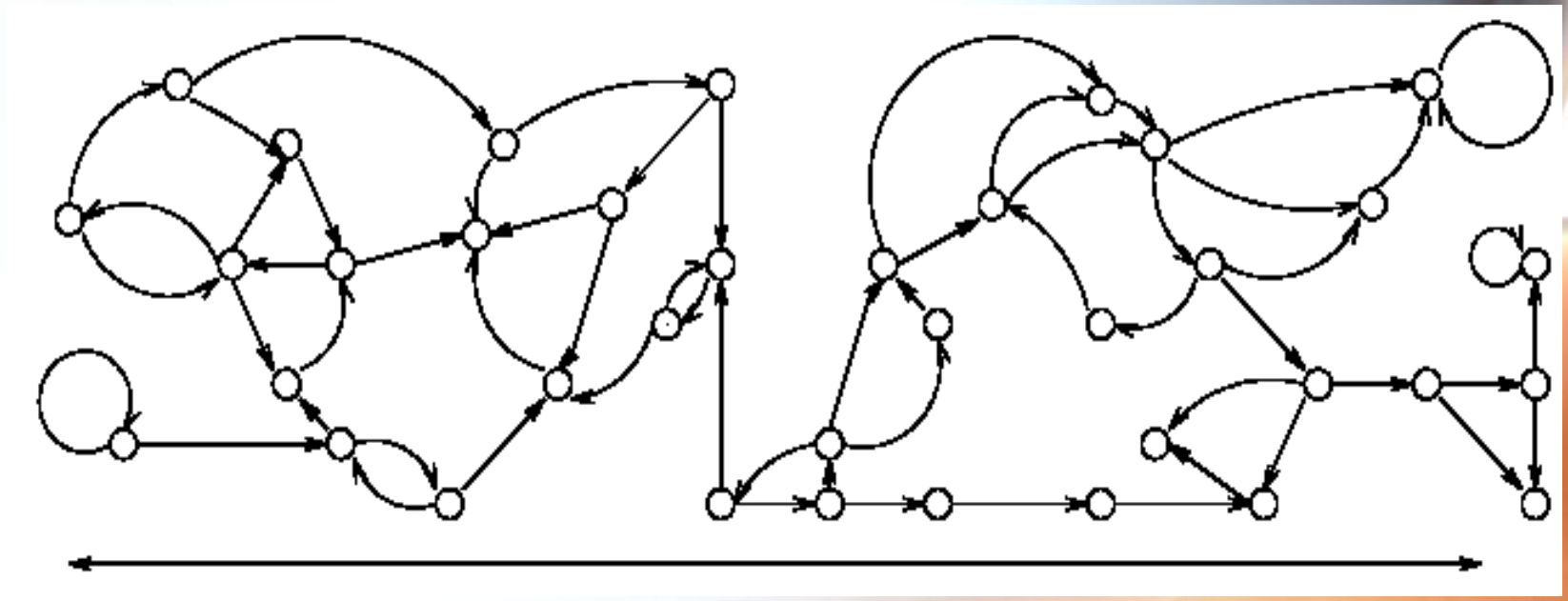
Small changes caused a gradual shift in thinking? Revolutions occurred gradually. Both theory and practical folks contributed to it.

Research Connectivity – Normal Science



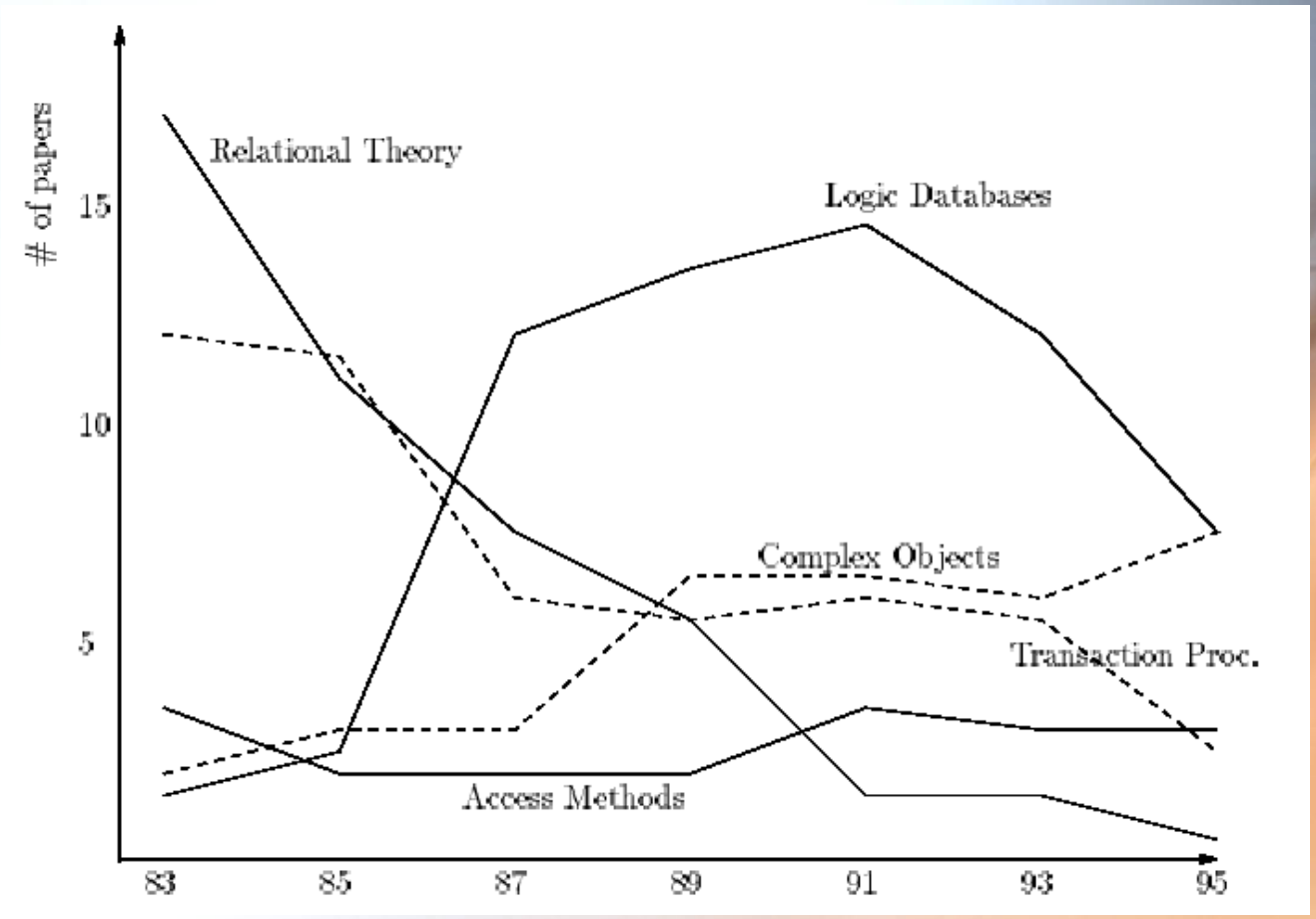
- Lots of connections
- Most theory within a few hops of practice, and vice-versa

Research Connectivity – Crisis



- Long paths from theory to practice
- Some nodes have no or little routes to practice
- In short term, this is very bad
- In *long* term, can help create new paradigm and new practice

Database Theory: a PODS retrospective



- 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

Christos's Theory Soapbox

- Good: Only now can one become a famous pure theoretician
- Bad: CS Theory is roundly bashed in some areas

And then there's applicability…

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

Final Discussion

It is not just theory but even applied computer science research is facing **low adoption rate** of ideas from the industry. There are new OS architectures, new Prog Languages, new ideas in DB, etc. Some ideas are rejected because they are incompatible with existing systems. Is this *backward compatibility* issue slowing our progress?

Conclusion

"I'll do anything to make people laugh. Just look at my hat!"

Christos Papadimitriou 5/25/08

UCB CS Commencement Ceremony