Doing Research

Systems for Interactive Data Exploration

Carsten Binnig
Brown University
Outline

1. What is a “good” contribution / paper?

2. How to write a good research paper?

3. Designing experiments?

4. Your paper for the class
Outline

1. What is a “good” contribution / paper?

2. How to write a good research paper?

3. Designing experiments?

4. Your paper for the class
Question 1: What is a “good” contribution / paper?

It is all about “taste”!
General Guideline

1 Paper = 1 Idea

No Buffet/Potpourri!
My Taste

Types of papers (ranked)
1. Vision / Innovative Systems papers
2. Experimental papers
3. Incremental / delta papers

Unfortunately: no papers about negative results
Type 1: Vision / Innovative Systems Paper

- It is all about the idea

- High potential to start a new line of research

- Many open questions is “good”

- Experiments to convince that vision makes sense but less important than for other papers
Type 3: Incremental Papers

• Tackles an *already known problem*

• **New method** to solve known problem
  – New algorithmic approach
  – New hardware ...

• **Benefits** of new method *must be crystal clear*

• Should **not leave many “open” questions**!

• **Experiments** should **show benefits / trade-offs**
Type 2: Experimental Papers

• Known Problem

• Known Techniques / Methods

• It is all about the experimental evaluation (should be main part of paper)

• Example: An Experimental Comparison of Thirteen Relational Equi-Joins in Main Memory
1. What is a “good” contribution / paper?

2. How to write a good research paper?

3. Designing experiments?

4. Your paper for the class
Question 2: How to write a good research paper?
Variant 1: Exploratory
Variant 2: Outline-driven
Today...

We are so frenetically pursuing the next conference deadline

...that the freedom required to do exploratory work is disappearing.

... and we often can not really finish the work (i.e., build a full system)
Three Factors Causing Problems

- Low acceptance rates
- Emphasis on paper count
- Bad reviewing
Outline

1. What is a “good” contribution / paper?

2. How to write a good research paper?

3. Designing experiments?

4. Your paper for the class
Question 3: Designing experiments?
Guideline 0: Be Reproducible

Hardware
- 56 nodes cluster
- CPUs / RAM / NICs
- Etc.

Software
- OS
- Language and compiler
- Etc.

Workload
- Queries
- Data
- Variations
ACM SIGMOD 2016 Reproducibility

July 6, 2016

All submissions should be directed to the DB Reproducibility email address. Click here to generate the email template. The sources for the code and the data will be communicated to the reviewers directly. The submission should accompanied by a detailed readme document based on this template.

July 5, 2016

New for this year is the Reproducibility Award which will be awarded to the 3 best papers. The awards will be presented during SIGMOD 2017 and each award comes with a financial prize of US$1500.

February 17, 2016

The SIGMOD 2015 papers that were awarded the reproducibility label are now available!
Guideline 1: First things first

Main claims should be shown first!

-> Use one “good” set of parameters
Guideline 2: Show other baselines

Optimal:

• Re-implement or
• use an-open source alternative

Also ok:

• If above does not work, compare to other papers
• use back-of-the envelope calculations
Guideline 2: Show other baselines
Guideline 3: Show other details / Vary parameters

Latency and Breakdown

Vary Degree of Distribution

Throughput (M trxs/sec)

Latency (us)

Probability of distribution

Cluster Size

Latency (us)

Cluster Size

Latency (us)

Probability of distribution
Outline

1. What is a “good” contribution / paper?

2. How to write a good research paper?

3. Designing experiments?

4. Your paper for the class
Question 4: Your paper for the class

Short write-up of 4-6 pages (not a research paper)

Use ACM templates: https://www.acm.org/publications/proceedings-template

Meta-outline of your paper
• Overview: Summarize your planned goals
• Architecture and System Details
• Demo Description (optimal: Link to Video)
• (Initial Experimental Results)
• Future Research Ideas (3 potential research avenues)