How to Write a Good Research Paper

Based on Simon Peyton Jones's Slides
Original by Liang Huang

September 16, 2014

Writing is NOT about English

Writing is not about language, but about logic

- Writing is equally hard for both native and non-native
- A bad paper is bad in any language

Different levels of writing

1. high-level(paper): global shape, logic, argument, style
2. mid-level(discourse): coherence with a paragraph
3. low-level(sentences): ordering of words and phrases
4. lowest-level(words): word choice, grammar

First Principle

Audience-Centric

- always have your audience or reader in mind
- writing is communication, NOT self-expression
- reader-centric attitude, not self-centric

Paper Communicate Ideas

- Infect the mind of your reader with your idea, like a virus
- Papers are far more durable than programs
- The greatest ideas are worthless if you keep them to yourself

Your Narrative Flow

- Here is a problem
- It's an interesting(important) problem
- It's an unsolved(hard) problem
- **Here is my idea**
- My idea works(details,data)
- Here's how my idea compares to other people's approaches

One Ping

- Read your paper again: can you hear the ping?

  One Ping one clear, sharp idea

- You may not know exactly what the ping is when you start writing, but you must know when you finish
- If you have lots of ideas, write lots of papers

Conference Paper Structure

- Title (1000 readers)
- **Abstract** (4 sentences, 100 readers)
- Introduction (1 page, 100 readers)
- The problem (1 page, 10 readers)
- My idea (2 pages, 10 readers)
- The details (5 pages, 3 readers)
- Related work (1-2 pages, 10 readers)
- Conclusions and further work (0.5 pages)

Abstract

- Used by program committee members to decide which papers to read, usually write the abstract last
- Four sentences [Kent Beck]
  1. State the problem
  2. Say why it's an interesting problem
  3. Say what your solution achieves
  4. Say what follows from your solution
- Example
  1. Many papers are badly written and hard to understand
  2. This is a pity, because their good ideas may go unappreciated
  3. Following simple guidelines can dramatically improve the quality of your papers
  4. Your work will be used more, and the feedback you get from others will in turn improve your research

Conference Paper Structure

- Title (1000 readers)
- Abstract (4 sentences, 100 readers)
- **Introduction** (1 page, 100 readers)
- The problem (1 page, 10 readers)
- My idea (2 pages, 10 readers)
- The details (5 pages, 3 readers)
- Related work (1-2 pages, 10 readers)
- Conclusions and further work (0.5 pages)

Introduction

- This is the hardest part of writing
  1. Describe the problem
  2. State your contributions
  3. ... and that is all
- Need to convey both importance and hardness

Method for Stating the Problem

- this is an **important** problem
- the dominant solution is good in A
- but bad in B (and B is important)
- the alternative solution is good in B but bad in A
- how to combine their merits?
- a **hard** problem!

State Your Contributions

- Write the list of contributions first, do not leave the reader to guess what your contributions are. The list drives the entire paper: the paper substantiates the claim you have made
- Reader thinks:

  If they can really deliver this, that's be exciting
  I'd better read on

- No rest of this paper is ..., Instead, use forward references from the narrative in the introduction. The Introduction (including the contributions) should survey the whole paper and therefore forward reference every important part.

Contributions should Be Refutable

- What does "refutable" mean? falsifiable and easily verifiable

	not refutable	refutable
	~~I'll devote myself to the American people.~~	I'll reduce unemployment rate by 5% by 2010.
	~~政府将尽全力为人民服务.~~	政府将在两年之内把PM 2.5降低50%.
you	~~our method is really powerful and elegant.~~	our algorithm is faster than the baseline by a factor of 1,000.

Conference Paper Structure

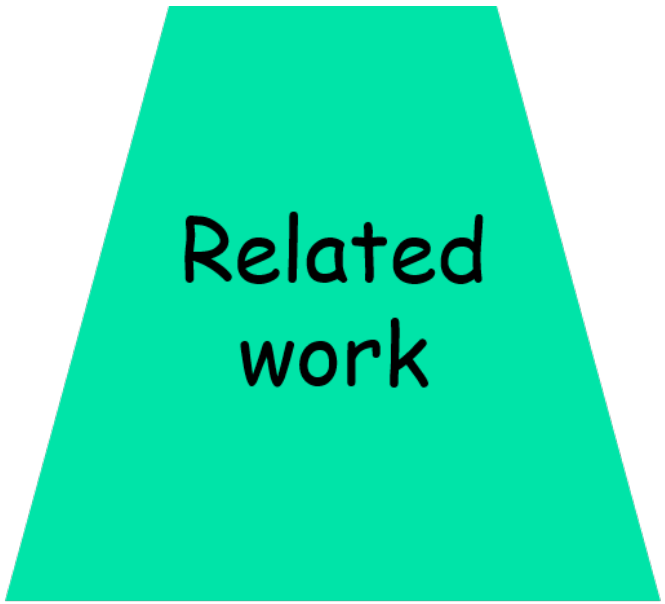
- Title (1000 readers)
- Abstract (4 sentences, 100 readers)
- Introduction (1 page, 100 readers)
- ~~**Related work**~~
- The problem (1 page, 10 readers)
- My idea (2 pages, 10 readers)
- The details (5 pages, 3 readers)
- Related work (1-2 pages, 10 readers)
- Conclusions and further work (0.5 pages)

No Related Work yet

Problem 1 carefully trimmed description of various technical tradeoffs is absolutely **incomprehensible** because reader knows nothing about the problem yet

Problem 2 describing alternative approaches gets between the reader and your idea

Your reader



Your idea

Conference Paper Structure

- Title (1000 readers)
- Abstract (4 sentences, 100 readers)
- Introduction (1 page, 100 readers)
- **The problem** (1 page, 10 readers)
- **My idea** (2 pages, 10 readers)
- **The details** (5 pages, 3 readers)
- Related work (1-2 pages, 10 readers)
- Conclusions and further work (0.5 pages)

Presenting the Idea

- Explain it as if you were speaking to someone using a white board. Conveying the **intuition** is primary, not secondary
- Once your reader has the intuition, she can follow the details but not vice versa. Even if skips the details, she still takes away something valuable
- Introduce the problem and your idea using **examples** and only then present the general case

The details

- Your introduction makes claims and the body of the paper provides **evidence** to support each claim
- Check each claim in the introduction, identify the evidence, and forward-reference it from the claim
- Evidence can be: analysis and comparison, theorems, measurements, case studies

Conference Paper Structure

- Title (1000 readers)
- Abstract (4 sentences, 100 readers)
- Introduction (1 page, 100 readers)
- The problem (1 page, 10 readers)
- My idea (2 pages, 10 readers)
- The details (5 pages, 3 readers)
- **Related work** (1-2 pages, 10 readers)
- Conclusions and further work (0.5 pages)

Related Work

Fallacy To make my work look good, I have to make other people's work look bad

Truth Credit is not loke money and giving credit to others does not diminish the credit you get from your paper

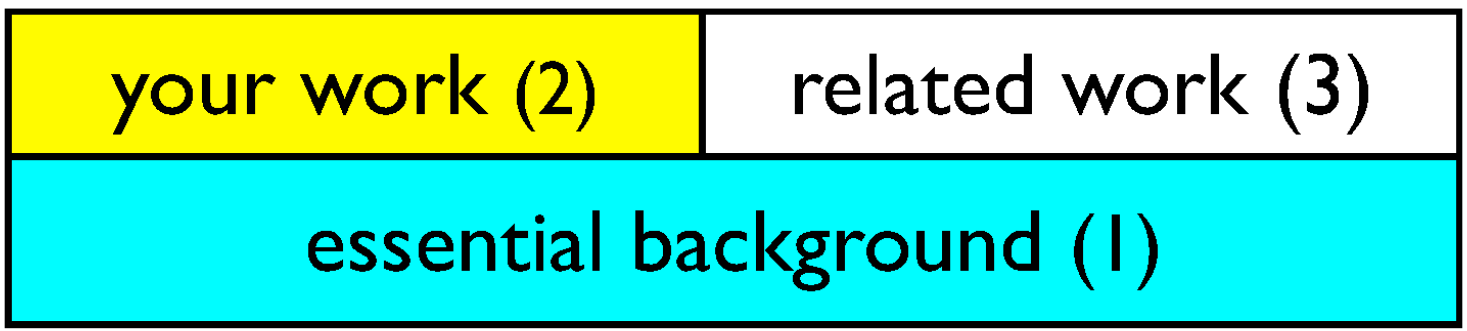
- Be generous to the competition
- Acknowledge weaknesses in your approach
- Failing to give credit to others can kill your paper.
- If you imply that an idea is yours, and the referee knows it is not, then either

  Bad You don't know that it's an old idea

  Worse You do know, but are pretending it's yours

Two Types of Previous Work

- essential background
  - the previous work that your work builds upon or improve upon
  - intro (w/o which readers can't understand your work)
- related work
  - other previous work that is just related to yours
  - having them doesn't change the understanding of your work



Conference Paper Structure

- Title (1000 readers)
- Abstract (4 sentences, 100 readers)
- Introduction (1 page, 100 readers)
- The problem (1 page, 10 readers)
- My idea (2 pages, 10 readers)
- The details (5 pages, 3 readers)
- Related work (1-2 pages, 10 readers)
- **Conclusions and further work** (0.5 pages)

Conclusions and Further Work

- Be brief.

Process of Writing

Start early, very early

- Hastily-written papers get rejected
- Papers are like wine: they need time to mature

Do not be intimidated

Fallacy You need to have a fantastic idea before you can write a paper. (Everyone else seems to)

Truth Write a paper and give a talk, about **any idea**, no matter how weedy it may seem to you

- Writing the paper is how you develop the idea in the first place

Listening to Reviewers

- Experts are good and non-experts are also very good
- Each reader can only read your paper for the first time once! So use them carefully
- Treat every review like gold dust and read every criticism as positive suggestion for something you could explain better
- Be truly grateful for criticism as well as praise because they have given up their time for you
- Do NOT respond "you stupid person, I meant X"
- Fix the paper so that X is apparent to the "stupidest" reader

Language and Style

- Remember to think of the paper as a collection of experimental results, summarized as clearly and economically as possible in figures, tables, equations and schemes.
- The text in the paper serves just to explain the data, and is secondary. The more information can be compressed into tables, equations, etc., the shorter and more readable the paper will be.
- Computer Science is NOT an experimental science, but you can still think of a paper as a collection of ideas, examples, algorithms, diagrams, definitions, theorems, proofs, plots and tables. Focus on the non-text parts and write text just to explain them.

Good Visual Structure

- Scientific evidences from *Paper Gestalt*
- A paper's fate(acceptance/rejection) can largely be determined by its visual features(layout) alone

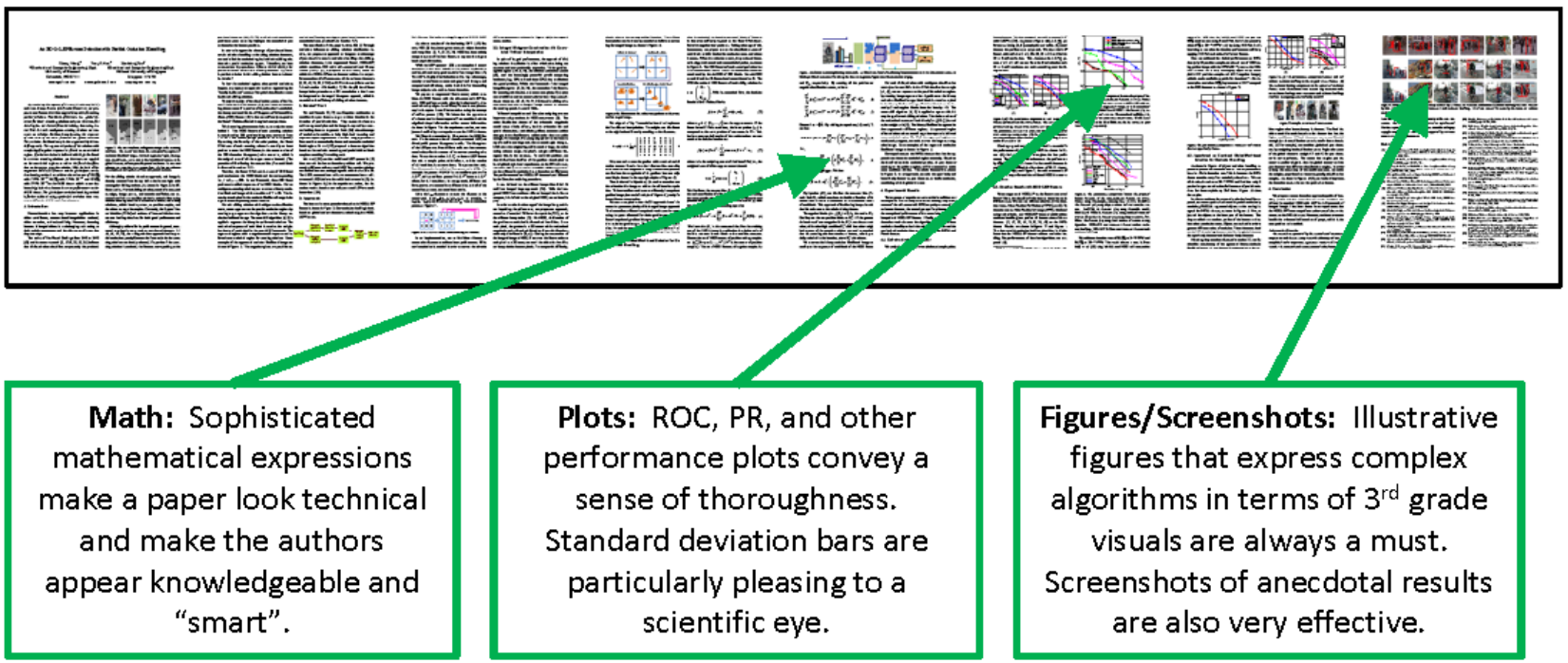


Figure 6. **Characteristics of a "Good" paper.**

Bad Visual Structure

- Scientific evidences from *Paper Gestalt*
- A paper's fate(acceptance/rejection) can largely be determined by its visual features(layout) alone

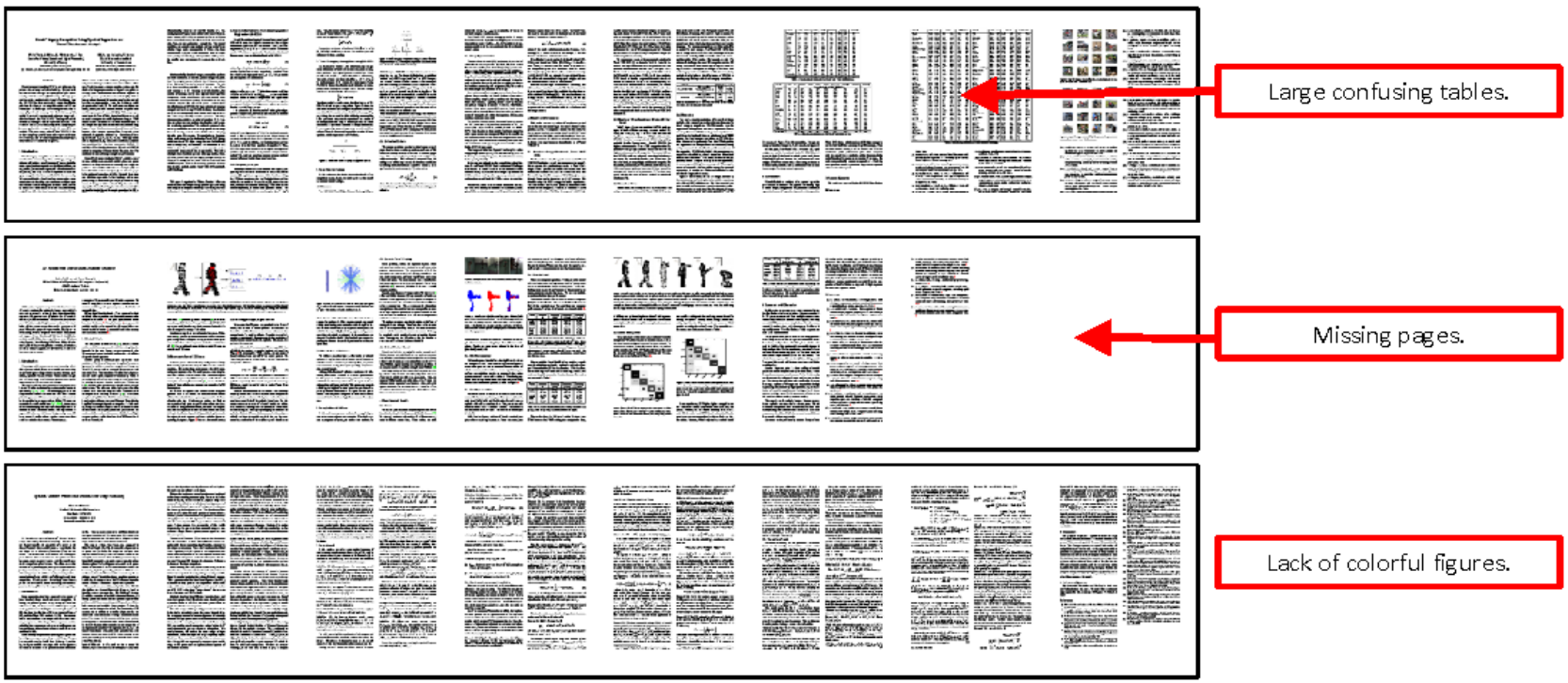


Figure 7. **Characteristics of a "Bad" paper.**

Resources for Scientific Writing

Fallacy students learn to write mainly from advisors

Truth learn from anybody whom you can learn from

High-Level

- Simon Peyton-Jones: *How to Write a Research Paper*
- Mark-Jan Nederhof: *Common Pitfalls in Academic Writing*

Low-Level

- Gopen&Swan: *The Science of Scientific Writing*
- Williams: *STYLE: Clarity and Grace series*
- Strunk and White: *The Elements of Style*
- Cook: *Line by Line*

Thank You

Q&A