How to write a Research Paper

Serge Autexier

German Research Centre for Artificial Intelligence Bremen (DFKI)

Doctoral Programme, CICM 2010
CNAM, Paris, France, 8. July 2010

Overview

- Why writing a paper at all (purpose)?
- How to organise a paper
- Kind of publication, when, where
- How does reviewing work?

Sources:

- Simon Peyton-Jones: How to Write A Great Research Paper
  `http://research.microsoft.com/en-us/um/people/simonpj/papers/giving-a-talk/writing-a-paper-slides.pdf`
- Alan Bundy: How to Write an Informatics Paper
  `http://homepages.inf.ed.ac.uk/bundy/how-tos/writingGuide.html`

Why Writing a Research Paper?

- Communicate Ideas, contribute to advancement of knowledge in your field
- Get recognition in your field (research career) to get promoted, get research positions
- People are less interested in technicalities of implementations (unless the topic is about programming in a specific programming language)
- They don't have your specific system, but want to get something reusable out of your work
- Sometimes ideas and implementation can get very close

How to Organise a Paper

There are different kinds of papers

- **Theoretical papers**: you have a problem (e.g. in Mathematics, theoretical Computer Science) and propose a solution
  - Is a problem decidable, semi-decidable?
  - What is the complexity of sorting a list ?
- **"Engineering" papers** (e.g. Computer Science, AI): you have a thesis, that can only be tested by experiments
  - Because the problem is not sufficiently explored to have a theory in which to study the question theoretically
  - Examples:
    - OCR for mathematical texts (Infty)
    - Automated Theorem Proving (in semi-decidable logics)
    - Daniel Kuehlwein's evaluation of premise selection for ATP
    - Melanie's approach to automate B set theory proofs by reduction to FOL

Important Parts of Papers

- Abstract
- Introduction
- The Problem
- Your Idea
- The details
  - Theory, Development, Theorems, Proofs
  - System Specification, Implementation, Evaluation
- Related Work
- Conclusion
- Future Work
- Appendices

The Title

- Ideally, the title should summarise the hypothesis of the paper.
- The reader should be able to work out what the paper is about from the title alone.
- Cute, cryptic titles are fun, but unhelpful.

  (James Davenport's suggestion: Use them as subtitles)

The Abstract

- The appetizer
- Also used by reviewers to select the papers they want to review
- Write it when the rest of the paper is written (or you have a clear structure)
- Must be self-contained and “closed”
  - No citations
  - No references into parts of the paper
- For instance in 4 sentences [Kent Beck]
  - State the problem
  - Why is this an interesting problem
  - What is your solution achieving
  - What follows from your solution

Introduction

- Brief context of your work
- Brief problem description
  - Maybe use an example to describe the problem (if adequate)
- and proposed solution (your contributions)
  - Write your contributions early to structure your paper
  - The later parts of the paper should substantiate your claims
  - Make it very explicit, for instance as bulleted list
- Use proposed solution to motivate/introduce theoretical bases you may need
- Introduce structure of your paper
  - either in text, for instance along with the contributions,
  - or as an explicit text

    The paper is organised as follows: In Section 2 we develop the foundations . . .

- At most 1 page

No Comparison to Related Work yet

- Problem 1: Comparing with related work before your idea gets between you and the reader
- Problem 2: It does not help the reader because she has yet nothing to check against

Reader gets tired

Reader loses interest

instead...

- Concentrate single-minded on a narrative that
  - Describes the problem
  - Describes you idea
  - Defends your idea, showing how it solves the problem
- On the way cite relevant work, but defer discussion/comparison to the end.

The Details: The payload of your paper

Consider a bifurcated semi-lattice D, over a hyper-modulated signature S. Suppose pi is an element of D. Then we know for every such pi there is an epi-modulus j such that pj ¡ pi.

- Sounds impressive. . . but
- Sends readers to sleep
- In a paper you MUST provide the details but FIRST convey the idea !!

Examples are crucial

- Examples are crucial to communicate ideas
- Must be well chosen
  - small and intuitive enough such that it can be easily introduced
  - it should allow to describe your problem and solution without being artificial
- Spend time to find and develop an example, look for “everyday” problems accessible to many readers.
- Use it to state the problem
- Use it to illustrate your technique
- Use it to illustrate how your technique works
- Introduce it as a running example very early in the paper (if not already in the introduction)

Conveying Your Idea

- Explain it as if you were speaking to someone using the white/blackboard
- Conveying the intuition is primary, not secondary
- Once the reader has the intuition, she can follow the details (but not vice versa)
- Even if she skips the details, she still takes away something usable

Evidence

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

Related Work

To make my work look good I have to make other peoples work bad

is a fallacy!

The Truth: Credit is not like money

Giving credits to others does not diminish the credit you get from your paper

- Warmly acknowledge people who have helped you
- Be generous for the competition.

  In his inspiring paper [Foo98] Foogle shows . . . We develop his foundation in the following ways. . .

- Acknowledge weaknesses in your approach as well as its limitations
- Honesty in science is essential and negative results are also important.
- Comparison of related work is part of the evaluation

The Truth: Credit is not like money

Sichere Kognitive Systeme

Not giving credits to others can kill your paper

If you imply that an idea is yours, and the referee/reader knows it is not, then either

- You don't know know that it's an old idea (bad)
- You do know, but are pretending it's yours (very bad)

Making Sure related work is accurate

- A good plan: when you think you are done, send the draft to the competition saying “could you help me ensure that I describe your work fairly?”
- Often they will respond with helpful critique
- They are likely to be your referees anyway, so getting their comments you front is good!

Conclusion & Future Work

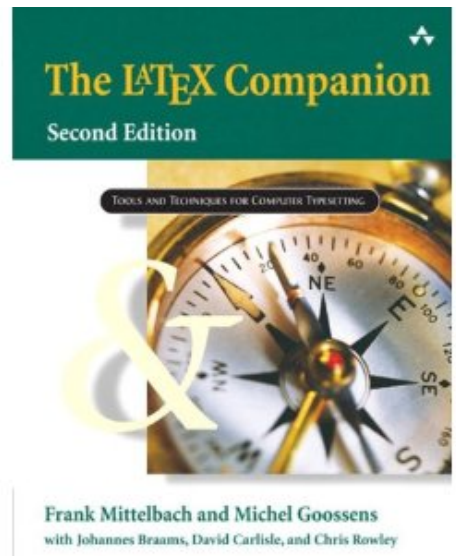
- More than a summary
- The conclusion should both summarise the research and discuss its significance
- Try to derive
  - what your solution shows
  - what can be learned from it
  - reassess the state of the field in the light of your contribution
- Future work:
  - Some unexplored avenues of the research
  - Identify and briefly develop new directions that have been suggested by your research

Appendices

- Use to provide additional material for the reviewing process to stay in page limits for main parts.
- Use it only if really, really necessary
  - it will not be in the final version, so the actual readers won't see it,
  - thus it should not be essential to understand the idea of your solution)
- For the final version: make a long version with all these details in a technical report and cite that one
- Additonnal material can be: proofs of less relevant lemmas, case study details, etc.

Bibliography

- Use a bibliography database the maintain and organise your bibliographic references
- Different kinds of publications (confence proceedings, journal articles, books, thesis) have different mandatory fields to fill
  Good description in Latex Companion 2nd Edition
- Build up while reading related work
- Use it for paper preparation (Bibtex format, but others exist)
- Nevertheless always check the generated bibliography at the end for duplicates, typos, etc.



Abstract, Introduction

- In **theory papers** you have claims substantiated by analysis and theorems (and their proofs)
- In **engineering paper** you must formulate a hypothesis and lay out by which methods you will evaluate it
- Not explicitly stating the hypotheses makes the contribution of papers vague
- Don't try to evaluate too many hypotheses at once, this makes the evaluation fuzzy and leads to confusion

Forms of Hypotheses

- Hypothesis can be of the following forms:

  (1) Technique/system X automates task Y for the first time;

  (2) Technique/system X automates task Y better, along some **dimension**, than each of its rivals;

- Dimensions:
  - **Behaviour**: X has a higher success rate than Y or produces better quality outputs
  - **Coverage**: X is applicable to a wider range of examples then Y.
  - **Efficiency**: X is faster or uses less space then Y.
  - **Dependability**: X is either more reliable, safe or secure than each of its rivals.
  - **Maintainability**: Developers find X easier to adapt and extend than its rivals.
  - **Useability**: Users find X easier to use than its rivals.

Specification & Implementation

- To conduct the evaluation, you need an implementation of your technique/system
- You should give a specification of your implementation, not only the description of your implementation (intuition vs. details)
- Specification:
  - The techniques that underlie the implementation are (formally) specified.
  - The requirements of the implementation are given.

Implementation

- Only the final state of the implementation should be described (not its history)
- The major design decisions should be identified and reasons given for the choices made.
- Abstract away from the code
- Outline the overall structure of the system and the key algorithms in abstract form,
  - e.g. using diagrams or formalised English/pseudo code.
  - A worked (running) example is often helpful.

Evaluation

- Evaluation is not testing
- Evaluation is the gathering of evidence to support or refute the hypothesis.
- Hypothesis 1: (first time):
  - system X must be applied to a sufficient range and diversity of examples of task Y to convince the reader that it constitutes a general solution to this task.
  - Descriptions of its behaviour, coverage and efficiency should be presented and, where appropriate, a description of dependability, maintainability or useability
- Hypothesis 2: (better, along some dimension, than each of its rivals)

  (Related Work!)
  - in addition to 1 there must also be a comparison with rival systems along the chosen dimensions
  - Also comparison along the unchosen dimensions, even if this is a negative result for system X;
  - honesty in science is essential and negative results are also important.

General Remarks

- Conveying an idea requires you guiding the reader
- No need to show how much you know about the whole area by writing an introduction to the whole field
- The reader is not that deep into the problem as you are
- Help the reader by explaining and avoid superflous details
- Be **concise** in order to not confuse the reader
- **Clarity/precision**:
  - Crucial for the reader and yourself
  - Unclear/obscure parts
    - Confuse the reader
    - Can give the reviewer the impression that something is odd/not well developed/not well understood by you
    - May indeed be parts you have not sufficiently understood/developed
  - Try to use short sentences in writing

General Remarks (cont'd)

- Quotations: When you quote other authors, give them the credit.

  "How to write a Research Paper"

  [Simon Peyton Jones]

  "How to write an Informatics Paper"

  [Alan Bundy]

- Quoting own previous work:
  - best rephrasing it than just copy and paste (context)
  - careful with reusing parts written together with collaborators (quotation)

Paper Writing Process

- **Start early, very early**
  - Hastily written papers get rejected
  - Papers are like wine: they need time to mature
  - To simplify improvements in terminology and notations use macro facilities
- **Proof read** your paper at least twice
  - Also proof read your bibliography
- **Collaborate** with other people and write papers with them
  - Collaboration helps finding good examples and explanations to get the idea through
  - Profit from experiences of seniors
- Use Version Control Systems (SVN) to collaborate
  Note: imposes format of the document you use for writing, not all document formats are well supported by VC systems, hence suitable for collaboration

Kinds of Research Publications

- **Technical reports**: institutions publication form
  - Good to start early with this to have a long and detailed description of all aspects of your work without page limitations
  - PhD Thesis (once out there) can play the same role
  - Can be freely re-used to write actual publication and also serve as a reference
- **Workshop papers**: possibly peer-reviewed, typically no publication, sometimes post-workshop proceedings or special issues in journals with new reviewing round
- **Conference papers**: peer-reviewed, real publications (Informatics and related), reputation depends on conference and publisher
- **Journal articles**: peer-reviewed, high quality publication (depends on Journal reputation and publisher)

Selecting the Publication Venue

- Search on the web for **conferences/journals** that have the topic of your work in the main topics
- Look where **related work** has been published
- Ask your **supervisor**
- Select publication venues that are **high quality**
  - publisher: IEEE, ACM, Springer, Elsevier, ...
  - Beware: Some major conference publish by themself (IJCAI)
  - referenced in major indexes like DBLP
  - Editorial boards (for journals)
  - Programm chairs (for conferences)
- Avoid **world conferences** and **multi-conferences** about everything and nothing
- **Check** the proceedings of previous events to get an idea of the **style of papers** are written in this venue
  - Different communities have different styles (mathematical, formal logic, technical vs. less technical, application oriented with e.g. UML, XML playing a major role)

Two slightly different submission and reviewing S

Sichere Kognitive Systeme

- Reviewed Workshops and conferences
- Journals (including special issues)

Conference / Workshop Reviewing

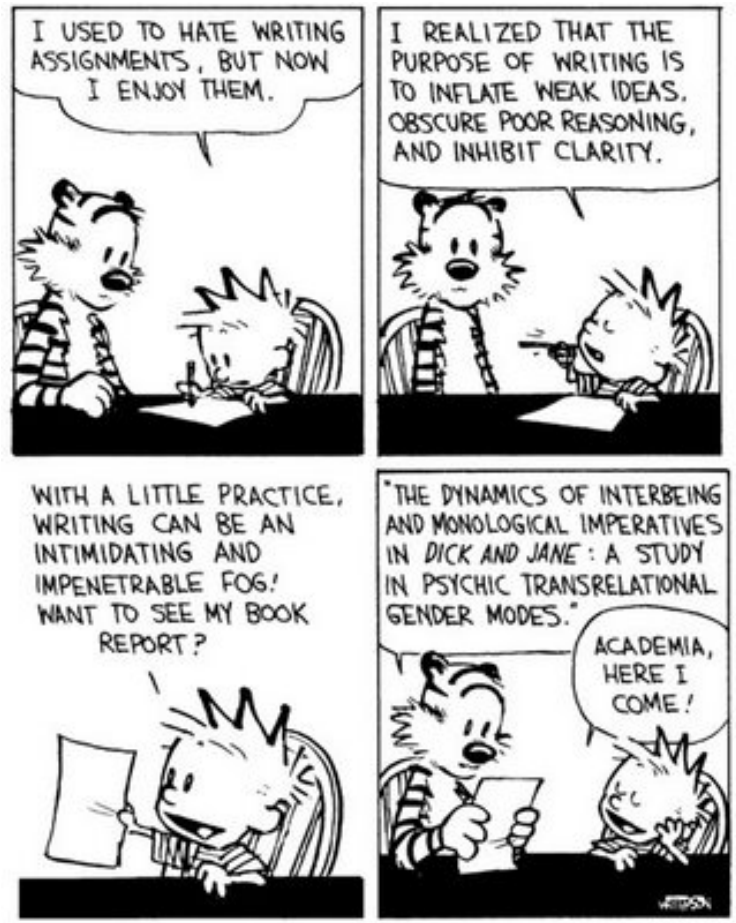
- **Submission**
  - Abstract submission
  - Paper submission
- **Paper assignment** to programme committee (PC) members by programme chair
  - Based on preferences indicated by PC members
  - Based on knowledge of the PC chair
- **Reviews** written by PC members, maybe by asking subreviewers
  - if it is not their area of expertise (selection often based on inspection of related work cited, but also by knowledge about who are the experts in the field)
  - to reduce their workload and get their students into writing reviews (to see the other side of the game)
- **PC discussion**: Reviews are discussed by PC members to come up with a decision of acceptance/rejection (or other forms of acceptance).
- Sometimes preceded by a rebuttal
- **Final decisions** made by PC chair

Journal Reviewing

- **Submission** to the editors (journal editors or guest editors in case of spec
- Paper assignment by editors based on their knowledge who the experts are
- **Reviews** written by expert reviewers
- **Reviews** are discussed among editors to come up with a decision of acceptance/rejection
  - Different option: major revisions and new round of reviewing

Summary & Pointers

Structure of Research Papers
Specific comments for Engineering Papers
General Remarks on Paper Writing
Kinds of Research Papers
Submission and Reviewing



- Alan Bundy: How to Write an Informatics Paper

  `homepages.inf.ed.ac.uk/bundy/how-tos/writingGuide.html`

- Simon Peyton-Jones: How to Write A Great Research Paper

  `research.microsoft.com/en-us/um/people/simonpj/papers/giving-a-talk/writing-a-paper-slides.pdf`

- Collected Advice on Research and Writing

  `www-2.cs.cmu.edu/afs/cs.cmu.edu/user/mleone/web/how-to.html`