Issues behind a research paper

Olivier Danvy

BRICS, University of Aarhus, Denmark

(danvy@brics.dk)

BRICS

KAIST, November 19, 1999.

A paper:
the prime medium for reporting scientific results.

In computer science: the conjunction of

- a good idea / a new idea,
- something sound, and
- something practical.

This talk: some issues behind a research paper.

Plan

- The variety of papers.
- Reading a paper.
- Refereeing a paper.
- Writing a paper.
- Targeting a submission.

The variety of papers

- Unpublished draft.
- Tech report.
- Workshop paper.
- Conference/symposium paper.
- Journal paper.
- And also: submitted / revised / final / extended.

The draft

Definition: the first shape assumed by a paper.

Use: for the author and for his immediate collaborators.

Quality control: the author(s).

The tech report

Definition: a draft readable by others.

Use: either as a time stamp (new draft with a new result),

or for the record (old draft with lots of detail).

Quality control: maybe a few colleagues & students.

The workshop paper

<u>Definition</u>: a record to document a talk.

<u>Use</u>: communication among specialists.

<u>Quality control</u>: the program committee (if any).

The conference/symposium paper

Definition: a record documenting a talk.

Use: communication in a larger community.

Quality control: the program committee.

The journal paper

Definition: the author's last word on a particular topic.

Use: archival purposes.

Quality control: the journal reviewers.

The tech report (revisited)

Definition: extended version of a conference / journal paper.

Use: typically includes all the proofs (and bypasses the copyright restrictions...).

Quality control: the reviewers.

The draft (revisited)

Definition: an unpublished paper.

Use: like wine, a draft may improve with age.

Quality control: unspecified.

Plan

- The variety of papers. $\surd$
- Reading a paper.
- Refereeing a paper.
- Writing a paper.
- Targeting a submission.

Reading a paper

- Information acquisition and retrieval.
- Critical reading.

Information acquisition

In principle, a paper provides enough information for its reader to reproduce its contents:

- proof,
- experiment.

Information acquisition

In principle, a paper provides enough information for its reader to reproduce its contents:

- proof,
- experiment.

Exercise: Think of the corresponding criterion

for an overview / survey paper.

Trust in the paper

Directly proportional to its advanced state:

journal versions are more trustworthy

than conference versions.

Information retrieval

It is a good idea to keep reading notes

(minimally as annotations in one's bibfile).

Critical reading

The three stages of reading (as one grows up):

1. the books say blah and thus it is true;
2. this book says blah;
3. the author wrote blah.

Examples

1. “Introduction to Data Bases”

2. “Advances in Data Bases”

3. “A new technique for query processing
   in object-oriented data bases”

More examples...

- “The λ-calculus with applications”
- “A λ-calculus notation for nameless dummies”
  (foreword by N. de Bruijn)
- etc.

Back on track

Here: paper, not book.

But one’s critical sense should still apply.

The alchemist's approach: "Solve et Coagula"

Don't swallow the author's propaganda,

accepting the paper as the author constructed it.

Instead:

- disassemble it to identify its real thrust;
- appreciate;
- probe / question / stress;
- (if needed) reassemble: minimize / expand.

Scientific reading (ended)

In principle, a paper provides enough information for its reader to reproduce its contents:

- proof,
- experiment.

But does it?

What if

- What if you don’t understand something in a paper?
- What if you think you found a bug in a paper?

Recommendation:

1. Consult people locally.
2. Send a $\boxed{\text{very polite}}$ e-mail to the author
   (keeping in mind that you may well be wrong yourself).

Plan

- The variety of papers. $\surd$
- Reading a paper. $\surd$
- Refereeing a paper.
- Writing a paper.
- Targeting a submission.

Refereeing a paper

<u>What</u>: the cornerstone of quality control.

<u>How</u>: peer review.

<u>Reference</u>: Parberry's guide for new referees.

Critical sense

One should not be impressed or unimpressed by the name and affiliation of the authors.

Example:

- D. Knuth, Stanford University.
- H. Potter, Hogwarts School.

—

Standard advice given to JFP referees:

It is a mistake to see depth where there is merely obscurity.

A few tips

- Reviews should be as $\boxed{\text{comprehensive}}$ as possible.
- Reviews should be as $\boxed{\text{courteous}}$ as possible.
- Reviews should be as $\boxed{\text{selfless}}$ as possible.

A concrete example of what can go wrong

David Harel's talk at DAIMI on Sept. 22, 1999.

First 10 minutes spent griping about anonymous referees, including the following quote from 1986 by a journal editor:

"I have personal antipathy towards languages based on hand-made graphics."

—

Analysis: not comprehensive, not courteous, not selfless.

Effect on the author: 13 years of griping!

More about refereeing

Parberry's guide for new referees.

Plan

- The variety of papers. $\surd$
- Reading a paper. $\surd$
- Refereeing a paper. $\surd$
- Writing a paper.
- Targeting a submission.

Writing a paper

A paper is written for others to read:

- yourself in the future, ☺
- other researchers, and
- reviewers.

It should thus reflect all the concerns already mentioned:

readability, clarity, etc.

Organization

- Title / list of authors / abstract.
- Introduction / compelling example / related work / overview.
- Development.
- Conclusion (if any).
- Acknowledgments / references.

The title

- It should be informative.
- It should be concise.
- It should be catchy / memorable.
- It needs to be original.
- It does not need to be funny.

On the temptation of being funny

The messenger can hide the message.

Circus analogy:

Do you want to be remembered as the clown (form)
or as the trapeze artist (contents / achievement).

Besides, most funny titles do not convey concrete messages: they tend to be puns or insiders's jokes.

The list of authors

- Alphabetically ordered;
- ordered by “degrees of contribution”;
- student first, adviser second;
- etc.

Key: the message should be more important
than [the order of] the messenger[s].

The abstract

- It should be brief.
- It should be as informative as possible.
- It should be updated last

  (to account for the actual contents of the paper).

The abstract

- It should be brief.
- It should be as informative as possible.
- It should be updated last
  (to account for the actual contents of the paper).

Fact:

Many more people will read your abstract than your paper
(e.g., in a bibliographic data base).

Titles, abstracts, data bases, and search engines

Abstracts are the key to locate papers on the web.

Abstracts are stored textually,

so they should NOT contain formulae, special symbols,

or bibrefs.

(Ditto for titles.)

As time goes by

It sounds silly, but...

- remember to date your manuscripts, and
- remember to update your bibrefs (“This paper is superseded by ...”).

The introduction

- It should start with a bang.
- It should stop with an overview of the rest of the paper.
- It should mention which prerequisites are expected from the reader.
- It should say clearly what is the achievement of the paper.
- A compelling example is always good, especially in a submission.

Starting with a bang

- “Real programs have effects.” (Mason and Talcott, JFP)
- “It was a dark and stormy night.” (Bulwer-Lytton)
- “Little Peggy was very careful with the eggs.” (Card)
- “The maid screamed, that was the problem with live servants.” (Goulart)

Eye catchers

- First words.
- Last words (of paragraph / section / chapter / thesis).
- Capitalization (in an abstract).

Pitfalls

- Exaggerating.
- Seeking effect for seeking effect:

  “This paper bridges a much needed gap.”

  (quoted by Neil Jones)
- Practicing Cooper’s prose style (cf. Mark Twain).
- Misspelling (always use a spell checker!).

Standing on the shoulders of giants

Disparaging earlier work invites the reader to disparage your own work.

Positivise – for example:

- the goal is blah, and it would be great to reach it;
- what has been done so far is remarkable,
  but it does not quite reach the goal;
- in this paper, we take one step further towards the goal.

Development

- Organized in sections.
- Should be progressive.
- Should be as complete as possible.
- Should be as concise and precise as possible.

Related work

- Mandatory.
- Situates the novelty and significance of your work.
- Where: either part of introduction, or part of conclusion, or stand-alone section (second or second to last).

Pitfalls: forgetting or mispresenting someone else's work.

Bibliographical references

Bibrefs should be used parenthetically, so that they do not interrupt one's reading.

Example: "...as seen in [2]." is awful, and so is
"[KAZAM97] shows that..."

Better: "...as introduced by Church in his monograph
on the λ-calculus [2]."

The conclusion

None in the mathematical tradition.

Minimally:

- recapitulates the problem and the contribution;
- assesses the significance of the contribution;
- suggests / outlines future work.

The pitfalls of one's future work

- Often presumptuous:

  "Writing a "future work" section of a paper is like a dog pissing on the trees at the boundary of its territory."

  (John McCarthy, 1991)

- Often shows the limits of the author's understanding.

The references

They $\boxed{\text{must}}$ be impeccable:

- accurate (correct year, etc.);
- complete (page numbers, etc.).

Standard pitfall: misspellings in titles.

Tips for writing a research paper

- In general, the contents should precede the paper. (But often, spelling things out in the paper tends to clarify its contents.)
- Top-down approach: goal and significance first.
- Bottom-up approach: results first.

A rule of thumb:

put in your paper what you (would) like to find

in other's papers.

Plan

- The variety of papers. $\surd$
- Reading a paper. $\surd$
- Refereeing a paper. $\surd$
- Writing a paper. $\surd$
- Targeting a submission.

A key resource

"How to have your abstract rejected"

Mary-Claire van Leunen and Richard Lipton

Conferences

Watch out for the theme of the conference.

Watch out for the program committee.

Watch out for what needs to be submitted:

- an extended abstract;
- a full version.

Extended abstracts (usually in the US)

The reviewers want to be convinced quickly.

The first impression is crucial:
abstract and introduction often determine the fate of the submission.

An extended abstract is closer to a workshop submission.

Full versions (usually in Europe)

The reviewers want the whole truth.

The simplest is to give it to them.

A full version is closer to a journal submission.

Classification (Parberry)

- Breakthrough.
- Ground breaking.
- Progress.
- Tinkering.
- Debugging.
- Survey.

Help the reviewer to make up his mind.

Methodological pitfall

Avoid core dumps.

- The paper should be focused on what it achieves.
- Tangents should be eliminated.

Motivational pitfalls

Ambitious: “Let’s write a paper for LICS.”

Opportunistic: “I’ve got to beef up my CV.”

Jealous: “I want more papers than X.”

Competitive: “I want to show that X’s papers are insignificant.”

Meteorological (esp. in Denmark): “Hmmm... Hawaii...”

etc.

Motivational pitfalls

Ambitious: “Let’s write a paper for LICS.”

Opportunistic: “I’ve got to beef up my CV.”

Jealous: “I want more papers than X.”

Competitive: “I want to show that X’s papers are insignificant.”

Meteorological (esp. in Denmark): “Hmmm... Hawaii...”

etc.

Better (if obvious): “I want to make science progress.”

Plan

- The variety of papers. $\surd$
- Reading a paper. $\surd$
- Refereeing a paper. $\surd$
- Writing a paper. $\surd$
- Targeting a submission. $\surd$

Summary and conclusion

A paper is part of research in computer science.

It is the means for reporting scientific results.

This talk:

how to write a paper and put it where it belongs,

ranging from a recycling bin to a library bookshelf.

Thank you.