

Writing a paper

Volodya Vovk

Input from: [John Shawe-Taylor](#) (now: UCL)

Department of Computer Science
Royal Holloway, University of London
Egham, Surrey, England

vovk@cs.rhul.ac.uk

29 November 2007

My plan:

- Scientists must write
- How to write a paper (content and form)
 - finding the information
 - overall structure and specific parts of the paper

How to publish a paper

A lot of lists; I know they are incomplete: please add.

Scientists must write

- A naturalist's life would be a happy one if he had only to observe and never to write. (Charles Darwin)

Scientists are different from other professions. Plumbers do not need to write about pipes. A scientific paper is the culmination of a scientific research.

How to write a paper

Standard scheme:

- think
- plan
- write
- revise (90%?)

Key Ingredient

Unless you are trying to bluff your way in science, a paper needs a **key ingredient**. John's table:

Key Ingredient	Risk	Payoff
Experimental testing or application of previously proposed technique	1	1
Review section of the literature	2	2
Generalization of previous result for another specialist area	2	2
New perspective on a result or results in one or more areas	3	4

Key Ingredient	Risk	Payoff
Sideways translation of previous result(s) into a new model or application	3	4
New algorithm for solving a known problem, perhaps more efficiently, or a novel heuristic for solving a difficult problem	4	4
Resolution of an open problem in a particular area	5	5
Novel theory to analyse some theoretical or real-world problem	5	6

How to arrive at the key ingredient?

- Start with an easy problem to get something under your belt.
- This will also get you familiar with the particular subject matter (the most important aspect of research is building up an intuitive picture of the area).
- Don't underestimate what you have: most new perspectives/results seem trivial with hindsight (this fact may be obscured by, e.g., complicated notation). **Don't talk to smart people about your best ideas while they are very new and delicate.**
- There is a natural progression from basic standard experimental research to novel ideas. Don't shy away from the mundane, but do some "venture capital investment" of time in trying to develop new ideas and perspectives.

Finding Information

Warning: you should not have too much respect for the printed word (especially if you are aiming at a high risk/payoff key ingredient).

It is that which we do know which is the greatest hindrance to our learning, not that which we do not know. (Claude Bernard)

The advantage of a certain amount of ignorance is that it keeps you from knowing why what you have just observed could not have happened. (Frederick Hopkins)

Reading before you think may prevent you from finding entirely new approaches to the problem.

But the final paper must cite and compare with relevant literature.

Sources of Information

- Internet
- Friends or supervisor
- Mailing lists
- Library
 - encyclopedias: general and specialized
 - special books and journals
(<http://www.rhul.ac.uk/information-services/library/>,
esp. <http://www.jstor.org/>)
 - Science Citation Index, <http://wok.mimas.ac.uk/> (follow
up a paper, yours or somebody else's)
 - abstracting journals

Structure of the Paper

Most of scientific papers: IMRAD (see Day's book).

- Introduction
- Methods (and Materials)
- Results
- Discussion

My usual scheme (suitable for theoretical papers): IRPAD.

- Introduction (why I have written this paper)
- Results and their interpretation
- Proofs
- Discussion

Other schemes?

Specific parts of a paper

Title: the most widely read part of the paper. The fewest possible words that adequately describe the contents of the paper.

Authors. Problems with who to include and the order of the names.

Table of contents: very useful but not always allowed.

Abstract: the map of your paper. See “title”; why you have written this paper.

Introduction: your purpose in writing this paper; background (such as existing literature); your principal results and conclusions. Remember the referees will probably be drawn from the references you cite.

“Discussion”: you can use it for discussion of weaknesses of your approach; interesting directions of further research.

Bibliography: no problems if using BibTeX.

Paper is not a story

About the early disclosure of your results (unlike O. Henry's endings):

Reading a scientific article isn't the same as reading a detective story. We want to know from the start that the butler did it.

Be clear in your own mind what is the take home message or slogan of your paper. Put it in the title and abstract and discuss/motivate in the introduction.

You won't be read unless you are accessible.

Is the form important?

If your paper is sloppy in form, 40% of your readers will not believe what you say! (Maybe subconsciously.)

If the paper contains a lot of maths, your readers will not be able to check everything and will need to rely on clues.

If somebody writes

“pp. 2-7” instead of “pp. 2–7”

you might think: he did not care to get his dots and dashes right; maybe the paper was written in a hurry and the argument is also flawed?

Minor fault(s) → major fault(s) suspected.

Clarity

Important factors of clarity:

- plain English (see Gowers)
- “readability”, as measured by different readability formulas
- grammar (active vs. passive voice etc.)
- proper use of \LaTeX

Readability formulas

Readability formulas used by Microsoft Word

After Word completes a grammar check, readability statistics are displayed, for example:

Flesch Reading Ease: This index computes readability based on the average number of syllables per word and the average number of words per sentence. Scores range from 0 (zero) to 100. The average writing score is approximately 60 to 70. The higher the score, the greater the number of people who can readily understand the document.

Limited usefulness: it is more important how the sentences fit together (does the next sentence require a double take?).

Ten commandments

In Day's book:

1. Each pronoun should agree with their antecedent.
2. Just between you and I, case is important.
3. A preposition is a poor word to end a sentence with.
4. Verbs has to agree with their subject.
5. Don't use no double negatives.

6. Remember to never split an infinitive.
7. When dangling, don't use participles.
8. Join clauses good, like a conjunction should.
9. Don't write a run-on sentence it is difficult when you got to punctuate it so it makes sense when the reader reads what you wrote.
10. About sentence fragments.

Don'ts

Things to avoid:

- starting a new paragraph after every displayed equation;
- forgetting to use “~” or “\ ” after a dot that does not end a sentence (initials etc.);
- writing variable x as x;

- forgetting to spell check the final version.

Thou shalt commit adultery. (Bible, published in England in 1631, the 7th commandment)

Know ye that the unrighteous shall inherit the Kingdom of God. (Bible, 1653)

Dos

Enhancing positive impression:

- paying attention to fonts;
- using different kinds of dashes for different purposes (–, —, —, —).

See almost any book on \LaTeX .

Publication process

What to do after your paper is written? **Oversimplification:** usually you write having some conference(s)/journal(s) in mind.

Choosing the right journal:

- opinion of your friends, colleagues, supervisor
- “objective” (easy to compute but measuring something else) measures: Journal Ratings in Science Citation Index

How to choose the right conference? Usually not so difficult. FOCS, STOC, IJCAI, UAI, ICML, COLT, Again a risk/payoff problem: you need to make an unbiased judgement of your paper.

Acceptance/rejection

Different reactions from the editor:

- acceptance (very rare)
- conditional acceptance (you should modify the paper)
- rejection

Rejection letter from a Chinese economics journal:

We have read your manuscript with boundless delight. If we were to publish your paper, it would be impossible for us to publish any work of a lower standard. And as it is unthinkable that, in the next thousand years, we shall see its equal, we are, to our regret, compelled to return your divine composition, and to beg you a thousand times to overlook our short sight and timidity.

How journals work

Peer review system.

Typical process for a journal submission:

1. You send your manuscript to the editor.
2. The editor (who is typically not paid) send your manuscript to several (typically 2 or 3) referees (your **peers**). The referees are given deadlines, usually a few months.
3. Referees (who are often busy people) study your manuscript and send their reviews to the editor.

4. The editor makes his/her decision based on the referees' comments; typically "reject" or "modify". You are informed about the decision. **After 6 months it is sensible to remind the editor about your paper.** "Reject": you send the paper to a different journal or forget the whole thing. Suppose the decision is "modify".
5. You decide whether you want to go on with your submission. If yes, you make those of the changes suggested by the referees that you agree with and explain in the cover letter to the editor why you disagree with the other changes. Go to step 1. In step 2 the editor is likely to send the paper to the same referees (or to accept your paper immediately).

Sometimes you send a manuscript to the editor-in-chief, who then sends the manuscript to an associate editor. Papers are usually accepted by the associate editor and rejected by the editor-in-chief (the best possible approximation to anonymity).

Even if a referee's remark is unfair, it is likely to point out to a problem: you should try to improve your exposition; a typical reader will have even less time than the referee.

How a conference works

Now I am interested only in rejection/acceptance of papers.

Whether your paper is rejected or accepted is decided by a programme committee. Different programme committees work differently.

A possible scheme (role of the Programme Chair small):

1. All members of the committee read some papers (say, at least 7) and assign to every paper read 2 numbers:
 - grade (e.g., 0 horrible, 10 excellent)
 - confidence (eg, 1 “not confident”, 3 “I am an expert in this field and this paper”)

The numbers and comments of all members are circulated to all other members by e-mail.

2. A heated discussion follows. In view of the discussion, a new set of 2 numbers is assigned. The Programme Chair asks some members to review papers which are reviewed by too few people.

3. The final discussion and the final numbers assigned.
4. Every submission gets a grade computed as the weighted average (the confidences serve as the weights) of the grades given to this submission.

The programme chair decides on the threshold; the papers above the threshold are accepted. Borderline papers: occasional exceptions.

Books

Robert A. Day, [How to write and publish a scientific paper](#), 3rd ed., Cambridge: Cambridge University Press, 1989. [501.49 DAY; K: 1998]

Newest incarnation: Robert A. Day and Barbara Gastel, [How to write and publish a scientific paper](#), 2006.

Robert Barrass, [Scientists must write](#), London: Chapman and Hall, 1978. [501.49 BAR]

Ernest Gowers, [The complete plain words](#), Penguin, 1987. [428 GOW]

Books on \LaTeX .

Books (cont.)

Great book of style:

[The Chicago Manual of Style](#), 15th ed., Chicago and London: University of Chicago Press, 2003.

First edition: 1906; 13th: 1982; 14th: 1993 [808.02 UNI].

Books recommended by Kostas Stathis:

Joseph M. Williams, [Style: Ten Lessons in Clarity and Grace](#), 8th edition, New York: Pearson, 2005.

Joseph M. Williams, [Style: Toward Clarity and Grace](#), Chicago and London: University of Chicago Press, 1990.

When writing a theory paper: see

Arieh Iserles, [How to write a paper](#), Arieh's web site (DAMTP, Cambridge), 2007.

+ the references therein