How to write a paper

Oded Goldreich
Department of Computer Science and Applied Mathematics
Weizmann Institute of Science, Rehovot, Israel
`oded.goldreich@weizmann.ac.il`

March 18, 2004

Abstract

The purpose of this essay is to make some self-evident (yet often ignored) remarks about how to write a paper. We argue that the key to writing well is full awareness of the role of papers in the scientific process and full implementation of the principles, derived from this awareness, in the writing process. We also provide some concrete suggestions.

Comment: The current essay is a revised and augmented version of our essay "How NOT to write a paper" (written in Spring 1991, revised Winter 1996). Originally, we confined ourselves to general principles and concrete negative comments. Here we overcame our reservation towards positive suggestions; hence the change in the essay's title.

Personal Comment: It is strange that I should write an essay about how to write a paper, because I consider myself quite a bad writer.[1] Still, it seems that nobody else is going to write such an essay, which I feel is in great need in light of the common writing quality in our community.

[1] My obsessive use of footnotes is merely one of the bad aspects of my bad writing style.

1 Introduction

This essay is intended to provide some guidelines to the art of writing papers. As we know of no way of cultivating artistic talents, we confine ourselves to some self-evident and mostly negative remarks. In addition, we make some concrete suggestions.

Our view is that writing a paper, like any other human activity, has some purpose. Hence, to perform this activity well, one has to understand the purpose of this activity. We believe that once a person becomes *totally aware* of his[2] goals in writing papers, the quality of the papers he writes (at least as far as form is concerned) will drastically improve. Hence, we believe that badly written papers are the product of either poor understanding of the role of papers in the scientific process or failure to implement this understanding in the actual writing process.

2 Why do we write papers or on the scientific process

The purpose of writing scientific papers is to communicate an idea (or set of ideas) to people who have the ability to either carry the idea even further or make other good use of it. It is believed that the communication of good ideas is the medium through which science progresses. Of course, very rarely can one be sure that his idea is good and that this idea may (even only eventually) lead to progress. Still in many cases one has some reasons to believe that his idea may be of value. Thus, the first thing to do before starting to write a paper is to ask *what is the idea* (or ideas) that the paper is intended to communicate. An idea can be a new way of looking at objects (e.g., a "model"), a new way of manipulating objects (i.e., a "technique"), or new facts concerning objects (i.e., "results"). If no such idea can be identified one should reconsider whether to write the paper at all. For the rest of this essay, we assume that the potential writer has identified an idea (or ideas) that he wishes to communicate to other people[3].

Having identified the key ideas in his work, the writer should first realize that the purpose of his paper is to provide the best possible presentation of these ideas to the *relevant community*. Identifying the relevant community is the second major step to be taken before starting to write. We believe that the relevant community includes not only of the experts working in the area, but also their current and future graduate students as well as current and future researchers that do not have a direct access to one of the experts[4]. We believe that it is best to write the paper taking one of these less fortunate people as a model of the potential reader. Thus, the reader can be assumed to be intelligent and have basic background in the field, but not more. A good example to keep in mind is that of a good student at the beginning stages of graduate studies[5].

Having identified the relevant community, we have to understand its needs. This community is undertaking the ambitious task of better understanding a fundamental aspect of life (in our case the notion of efficient computation). Achieving better understanding requires having relevant information and rearranging it in new ways. Much credit is justifiably given to the rearrangement of information (a process which requires "insight", "creativity" and sometimes even "ingenuity"). Yet,

[2]For simplicity we chose to adopt the masculine form.

[3]We leave the case of criminals that pollute the environment with papers in which even they can identify no ideas, to a different essay...

[4]Indeed the chances that the experts (in the area) will be the ones that further develop or use the new ideas are the greatest. Yet, much progress is obtained by graduate students and/or researchers who became experts only after encountering these new ideas and further developing or using them.

[5]Ironically, the writers who tend to care the least about readers that are at this stage of their development (i.e. beginning of graduate school) are those who have just moved out of this stage. We urge these writers to try to imagine the difficulties they would have had if they had tried to read the paper, just being written, half a year ago...

the evident importance of having access to relevant information is not always fully appreciated[6]. The task of gathering relevant information is being constantly frustrated by the disproportion between the flood of information and the little time available to sorting it out. Our conclusion is that it is the writer's duty to do his best to help the potential readers extract the relevant information from his paper. The writer should spend much time in writing the paper so that the potential readers can spend much less time in the process of extracting the information *relevant to them* out of the paper.[7]

3 How to serve the reader's needs

In the previous section, we presented our belief that the purpose of writing a paper is to communicate a set of ideas to researchers that may find them useful. As these people are drowning in a flood of mostly irrelevant information, it is extremely important to single out clearly the new ideas presented in the paper. Having understood the abstract requirements, it is left to carry out this understanding to each level of the writing process: from the overall structure of the paper, through the structure of single paragraphs and sentences, to the choice of phrases, terms and notation. Here are some principles which may be useful.

3.1 Focusing on the readers' needs rather than on the writer's desires

The first part of the above title seems mute at this point, yet the second part warns against an evasive danger that may foil all good intentions: The writer is often overwhelmed by his own desires to say certain things and neglects to ask himself what are the real needs of the reader. The following symptoms seem related to the latter state of mind.

- *The "Checklist" Phenomenon*: The writer wishes to put in the paper everything he knows about the subject matter. Furthermore, he inserts his insights in the *first* possible location rather than in the *most* suitable one. In extreme cases, the writer has a list of things he wants to say and his *only concern* is that they are *all* said *somewhere* in the paper. Clearly, such a writer has forgotten the reader.

- *Obscure Generality*: The writer chooses to present his ideas in the most general form instead of in the most natural (or easy to understand) one. Utmost generality is indeed a virtue in some cases, but even in these cases one should consider whether it is not preferable to present a meaningful special case first. It is often preferable to postpone the more general statement, and prove it by a modification of the basic ideas (which may be presented in the context of such a special case).

- *Idiosyncrasies*: Some writers tend to use terms, phrases and notations that only have a personal appeal (e.g., some Israelis use notations which are shorthand for Hebrew terms...). Refrain from using terms, phrases or notations that are not likely to be meaningful to the

[6]Of course, everyone understand that it is important for him to have access to relevant information, but very few people care enough about supplying the community with it. Namely, most people are willing to invest much more effort in obtaining a result than in communicating it. We believe that this tendency reflects a misunderstanding of the scientific process.

[7]This imperative is justified not merely by abstract moral reasons, but rather out of practical considerations related to the economy of resources. Firstly, the number of people that would read the paper is typically significantly larger than the number of its authors. Secondly, typically, the authors have much better understanding of the subject of the paper and it should take them less time to figure out missing details or articulate the ideas.

reader. The justification to using a particular term, phrase or notation should be its *appeal* to the intuition or the associations *of the reader.*

- *Lack of Hierarchy/Structure*: Some people can maintain and manipulate their *own ideas* without keeping them within a hierarchy/structure. But is it very rare to find a person who will not benefit from having new ideas presented to him in a structured/hierarchical manner. Specifically, the write-up should make *clear distinctions* between the more important ideas/statements and the less important ones. That is, one should highlight important ideas/statements, and mark secondary discussions as such. The specific ways of "highlighting" and "marking" may vary, but they should be conspicuous.

- "*Talmud-ism*": The writer explores all the subtleties and refinements of his ideas when *first* introducing the and before clarifying the basic ideas. Furthermore, the writer discusses all possible criticisms (and answers them), before providing a clear presentation of the basic ideas.

All these symptoms are an indication that the writer is neglecting the readers and their needs, and is instead concentrated in satisfying his own needs.

3.2 Awareness to the knowledge level of the reader

Another difficulty involved in the process of writing is lack of constant awareness to what the reader may be expected to know at a particular point in the paper. Some points to consider are:

- Whenever presenting a complex concept/definition, beware that the reader cannot be assumed to fully grasp the new concept and all its implications immediately.

- Whenever presenting proofs be sure to *elaborate* on the conceptual steps rather than on the standard technical analysis. Having done the conceptual steps yourself, they seem rather evident to you, but they may not be evident to the reader. Furthermore, these conceptual steps are typically the most important ideas in the paper and the ones with which the readers have most difficulties.

- As said above, one should try to avoid treating the general case with all its complications in one shot. Thus, one may first present a special case that captures the main ideas and later derive more general statements by introducing additional (secondary) ideas. Whenever this is done, try to obtain the general results by either use of reductions to the special case, or by high level modifications to it. Try to avoid the use of syntactical (or local) technical modifications of the special case as a way to obtain the general case.

- Don't hide a fundamental difficulty by using a definition that ignores it without first discussing the issue (i.e., what is the difficulty and why bypassing it does not deem the entire investigation meaningless).

- Try to minimize the amount of new concepts and definitions you present. The reader's capacity of absorbing concepts and definitions is bounded.

3.3 Balancing between contradictory requirements

The suggestions made in the above subsections may be contradictory in some cases. Such cases call for the application of judgment. The problem is to balance between contradicting requirements. Indeed this is a difficult task.

Application of judgment requires *flexibility*. The writer should NOT try to follow a canonical example or structure, but rather apply good principles to the concrete problems and dilemmas emerging in writing the current paper.

3.4 Making reading a non-painful experience

Following are some common examples of writing mistakes which make reading a very painful experience:

- *A labyrinth of implicit pointers*: The words "it" and "this" are commonly used as implicit pointers to entities mentioned in previous sentences, but the reader can find it difficult to figure out to which entities the writer was referring. Consider, for example, the following sentences "A is interested in doing X. It has property Y but not Z. This property allows it to do this". The writer should consider making these pointers explicit (by explicitly referring to objects by their names).

- *Sentences with complex logical structure*: Technical papers introduce a host of specific parsing problems. One type of problems is introduced by sentences with complicated logical structure (i.e., conditional sentences, having multiple and sometimes nested conditions and consequences, like "if X and Y or Z then P or Q").

- *Mixture of mathematical symbols and text*: Consider, for example the sentences "on input x, y, A runs B^y on $f(x)$". A more clear alternative is "on input x and y, algorithm A runs the oracle machine B on input $f(x)$ placing y on B's oracle tape". It never hurts reminding the reader of the categorical status of the objects.

- Cumbersome notations and terms: For example, consider an object denoted $M_{i_{j,k_t}}^{O_b^c}$, or a multiple parameters term like an $(a, b, c, d, e, f, g, h, i, j)$-system or a multiple qualifications term like a *kuku-muku popo-toto system.*

4 Some concrete suggestions

As hinted above, we do not believe that there is one good format that should be followed in all papers. On the contrary, a key ingredient in the process of writing is flexibility: the selection of a form that fits the contents at hand. This selection requires the application of judgment in order to balance between various contradicting concerns.

The aforementioned beliefs were the reason for our original decision not to provide any concrete positive suggestions. However, in retrospect we realized that it may be valuable to make some concrete positive suggestions while warning that these are not of absolute nature and are only useful in most cases (rather than in all cases).

Indeed, there are some general rules that apply to almost all papers. Typically, a paper must have a title, a list of authors, an abstract, an introduction, a main part, and references. We will discuss these components and related issues next.

4.1 The title and the list of authors

The title should be as informative as possible and yet not too cumbersome or too long. Indeed, one cannot fit much information in a title, but one better provide some clue about (or at least hint to) the contents of the paper (rather than make jokes). One should bear in mind that the paper's title should fit into a sequence of past and future work: Hence, the title should be sufficiently different from the titles of previous works (and should allow room for subsequent works).

The common tradition in our community is to list authors in alphabetical order. This tradition seems linked to the typical situation in which each author has made a fair (if not significant) contribution to the work.[8] Since this is the common practice in our field, deviating from it in special cases does not make much sense, and we strongly recommend not to do so but rather find alternative ways to compensate for vast inequality in contribution to the work.[9] We note that failure to follow this suggestion may cause more harm than good to the person pushed forward in the alphabetical order (e.g., it may encourage various committees to ask why this person is not the first author in other papers, etc).

4.2 The abstract

The abstract should be as informative as possible and yet not too cumbersome or too long. Indeed, the same was said of the title, but in case of the abstract the space allocation is ten-fold or twenty-fold increased. Still, typically, one cannot (and should not) provide a rigorous definition and detailed statement of results in the abstract. Instead, one should provide a high-level description of the contents of the paper. One should bear in mind that, for a variety of reasons, some people will only read the abstract and one should provide these people with as much information as possible. Furthermore, some people may access the paper in ways that do not allow them to look at other parts of the paper (and so it is strongly recommended (and sometimes even required) that the abstract should not refer to other parts of the paper (e.g., to the list of references)). That is, the abstract should be self-contained. On the other hand, the abstract should not be long (because then it stops being an abstract). Typically, the abstract should not exceed 200 words.

There is a clear contradiction between the desire that the abstract be self-contained and the impossibility of making it *really* self-contained. But the abstract should not be really self-contained. *It should be self-contained only as a high-level description of the contents of the paper.* This is possible because:

- The abstract need not motivate the model (as this will be typically done in the introduction).
- The abstract need not list and/or recall the contents of prior work (but rather, if necessary, it may describe the nature of the improvement over possibly unspecified "prior work").
- The abstract need not provide an accurate description of the paper's results (but may rather describe them in imprecise but clear terms using warning phrases as "loosely speaking"). In cases where even an imprecise (but clear) description is infeasible, the abstract may merely convey the flavor or nature of the new results.

[8] Indeed, non-alphabetical order is common in disciplines in which the common situation is different (e.g., a lab-head co-authors any work done in his/her lab, and various levels of technical assistance are acknowledged by listing these helpers as co-authors).

[9] For example, if one person has made a negligible contribution to the paper then this person better retire from it and get acknowledged (in the Acknowledgments) rather than be made the last co-author (which may be ineffective in case his/her name is anyhow alphabetically last).

- The abstract need not provide a description of all the paper's results (but may rather confine itself to the most important ones, while clarifying that these are merely the main results).

Note that we are not saying that the abstract *should not do* any of the aforementioned things, but rather that it *does not have to do them*. Indeed, making a choice of what to provide in the abstract, calls for exercising judgment (based on deep understanding of the work).

We stress again that one should bear in mind that the abstract is all that may be available to some readers. This was true even before the days of the WWW (e.g., collections of abstracts or digests of them were popular in the past), but is certainly true nowadays (when some web-services either provide access only to the abstracts of works or operate based only on such abstracts).

4.3 The introduction

Typically, the introduction should provide a clear description of the work as well a good motivation to it and a comparison to prior works. That is:

- The introduction should provide a clear description of the contents of the paper. In particular, the introduction should provide a clear statement of the main results and a high-level description of the techniques. The level of detail of these descriptions may vary: In most cases it should be possible to provide sketchy versions of the main theorems and to describe the main ideas underlying the techniques, but this is not always possible. In the latter cases, an adequate alternative should be found.

  The introduction should highlight important new ideas and novel conceptual observations. In case it is not feasible to describe these elements without the technical context, the introduction should state their existence and refer the reader to the place in the paper were they can be found.

- Typically, the introduction should provide a clear motivation to the questions studied in the paper. Exceptional cases refers to well-known questions having well-known motivation. Assuming that the readers know the motivation is a calculated risk, but sometimes such risks are worthwhile taking. The motivation need not be argued from scratch. If there are dozens of works dealing with a particular type of questions then this type requires no motivation, but the specific question within this type may require motivation.

  Regarding work in the theory of computation, my own opinion is that a good motivation is one that connects the current study to central notions and questions of the relevant area. The connection should be natural (i.e., not contrived), and it should make sense with respect to the actual study (rather than be only falsely related to it). There is no need to provide an "actual application" (although a good one may demonstrate the viability of the connection).

- The introduction should place the current paper in context of prior related work. Assuming such related work exists, the (main) differences with respect to it should be pointed out and fairly evaluated: The aspects in which the new paper improves over prior work as well as aspects in which it is worse should be clearly stated and discussed.

Good ways of providing the aforementioned information may vary from paper to paper (and there is no "best" way). There is no specific order in which one should proceed, nor any canonical pattern to follow. Needless to say, there is no universal level of detail that should be provided. One should select a structure that makes sense, and follow it with care (keeping the reader in mind). The final test is the reader: did he/she obtain a good idea about the contents of the paper?

Important conclusions and natural open problems that arise from the current work (rather than well-known ones) may also be stated in the introduction. Stating these elements in the introduction is preferable to stating them them in a conclusion section, unless these elements are significantly easier to understand after reading the main part of the paper. See related discussion in Section 4.5.

4.4 The technical part

There is little we can say regarding the main part of the paper, beyond the general principles outlined in Sections 2 and 3. Still, we highlight a few concrete implications of these general principles:

- *Discussing definitional choices*: Definitions embody a host of decisions ranging from the choice of the notion that the definitions intend to capture to very technical choices (which may be either arbitrary or important). It is important to provide insightful discussions of these definitional choices. This holds not only for the high-level choices but also for low-level choices. High-level choices should be motivated by linking them to the notions that the paper studies (which were already motivated in the introduction): The connection may be obvious (in which case a reference suffices), but otherwise a good discussion is called for. Low-level choices may be even more problematic. It is important to say whether these choices are *arbitrary*, *simplifying* or *essential* (in the following sense):

  – A choice is arbitrary if almost any other reasonable choice will have the same effect.

  – A choice is adopted for sake of simplicity if it can be replaced by more natural choices at the cost of (merely) complicating the discussion.

  – A choice is essential simplicity if it is essential to the claimed results, which are not *known* to hold when adopting an alternative choice that seems as natural (or as reasonable).

  The latter case is indeed disturbing, especially if one cannot provide a good explanation as to why these seemingly technical choices are important. Still one should be honest about it.

- *On numbering technical elements*: Unless the paper is very short (e.g., less than five pages long), it is important to use a numbering system that supports easy searches for a given item. Indeed, it is "logical" to use a different counting-number for each type of element (e.g., Definition 5 would be the fifth definition in the paper, Theorem 3 the third theorem, and so on), but this traditional convention makes finding a specific element quite hard. The alternative we advocate is using a single numbering system such that items can be easily found by binary search... (In case of long paper, we suggest using a double-numbering system by which Theorem 5.2 is the second item in Section 5).

4.5 Conclusions and/or suggestions for further work are not a "must"

Some people tend to think that each paper should end with conclusions and/or suggestions for further work. We strongly disagree with this opinion, and see little use in a "conclusion section" that merely re-iterates things said in the abstract and/or in the introduction. Similarly, we see no point in listing well-known open problems or re-iterating questions that were already raised in the introduction. On the other hand, we do value a conclusion section that contains high-level material that better fits *after* the main part of the paper (and thus is not placed in the introduction). Similarly, for raising important questions that are more appealing after reading the technical part (even if they were raised already in the technical part but not in the introduction).

To summarize: There are papers that may benefit from a conclusion section, but they are relatively few (say, less than 5% of the papers). Certainly, the inclusion of a conclusion section should not be the default.

4.6 References and acknowledgments

A delicate issue that comes up when writing a paper is that of referring to other works and acknowledging help from other researchers. Two (sometimes contradictory) principles that may govern our decision are truth and kindness. As our primary concern is providing information, truth is of utmost importance. We should never mislead the reader by unjustified or inaccurate credits attributed to other works. But within the domain of truth one should be kind. For example, the reader will not be harmed if the writer acknowledges each person with whom he had a relevant discussion.

5 Benefiting from readers' comments

Occasionally, writers ask their friends and close colleagues for comments on their write-up. Typically, these comments are not useful because friends and close colleagues feel reluctant to point out major expositional problems. Furthermore, these friends and close colleagues may know the work before reading it or at least may have a better a priori knowledge about the work than an average reader may have. In any case, it is very dangerous to conclude from the fact that the writer's friend (or close colleague) liked the write-up that the write-up is indeed good. (Needless to say, it is dangerous to conclude from the fact that the writer likes the write-up that the write-up is indeed good.)

Thus, if you ask a friend (or close colleague) to give you comments, make sure this friend understand that you are interested in a critical reading and not in compliments.

Readers that may be assumed to be critical are reviewers. They typically point out problems and make suggestions. One should *not necessarily* follow the reviewer's suggestions, but one must always bear in mind that theses suggestions indicate problems in the current write-up. It may be that the reviewer is not suggesting a good solution to these problems (or that the authors has a better a solution), but for sure there is a problem.[10] That is, the working assumption (which is almost always correct) is that any comment of a reviewer indicates a problem in the write-up: Reviewers are typically not idiots, and one can learn even from idiots!

[10]Needless to say, if the author decides not to adopt a reviewer's suggestion in a the course of a review process for a journal publication, then the author should justify this decision in a letter to the handling editor.