How NOT to write a paper*

Oded Goldreich
Department of Computer Science and Applied Mathematics
Weizmann Institute of Science, Rehovot, ISRAEL
Email: oded@wisdom.weizmann.ac.il

December 21, 1996

Abstract

The purpose of this article 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.

1 Introduction

This article is intended to provide some guidelines to the art of writing papers. As there can be no way of cultivating artistic talents, we must confine ourselves to some self-evident and mostly negative remarks.

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[1] 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 either of poor understanding of the role of papers in the scientific process or of lack of implementation of this understanding.

2 Why do we write papers, or the Scientific Process

The purpose of writing scientific papers is to communicate an idea (or ideas) to people who have the ability to carry the idea even further or to 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. "model"), a new way of manipulating objects (i.e. "technique"), or new facts concerning objects (i.e., "results"). If no such idea can be identified one should reconsider writing a paper at all. For

*Written in Spring 1991, revised Winter 1996
[1]For simplicity we chose to adopt the masculine form.

the rest of this article, we assume that the potential writer has identified an idea (or ideas) which he wishes to communicate to other people[2].

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 consists not only of the experts working in the area, but also of their current and future graduate students as well as of current and future researchers which do not have a direct access to one of the experts[3]. We believe that it is best to write the paper taking one of these less lucky persons as a model of the potential reader. Thus, the reader can be assumed to be intelligent and have basic background in the field, but no more. A good example to keep in mind is that of a good student at the beginning of graduate school[4].

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 a new way. Much credit is justifiably given to the rearrangement of information (a process which requires "insight", "creativity" and sometimes even "ingenuity"). Yet, the evident importance of having access to relevant information is not always fully appreciated[5]. The task of gathering relevant information is being constantly frustrated by the disproportion between the flood of information and the little time available to sort 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.

3 On Implementing the above Principles

In the previous section, we presented our belief that the purpose of writing a paper is to communicate a set of ideas to researchers which 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.

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

[3]Indeed the chances that the experts (in the area) will be the ones which 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.

[4]Ironically, the writers who tend to care the least about readers which 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....

[5]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.

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 which may foil all good intentions: The writer is often governed by his own desires to say certain things and neglects to ask himself what are the real needs of the reader. The following symptoms seems 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* suitable location and not 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.

- *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 (presented within a special case).

- *Idiosyncrasies*: some writers tend to use terms, phrases and notations which have only a personal appeal (e.g., some Israelis use notations which are shorthand for Hebrew terms...). Refrain from using terms, phrases or notations which are not likely to be meaningful to the 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 his in a structured/hierarchical manner. Specifically, the write-up should make *clear distinctions* between the more important ideas/statements and the less important ones.

- *Talmudism*: the writer explores all the subtleties and refinements of his ideas when *first* introducing them. He discusses all possible criticisms, answers to these criticisms, and so on, before the reader had a chance to get 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 shoot. Thus, one may first present a special case which 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 which 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 a case calls 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 to do X. It has property Y but not Z. This property allows it to 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 vast 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, 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.

4 References and Acknowledgments

A delicate issue which 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 by an overly excessive acknowledgment to somebody or by a generic reference to some work as part of a list of previous relevant works.