Plagiarism and Other Sins

In the world of academia, art, science, and any discipline whose foundations are based on creative thinking and intellectual property, plagiarism is arguably one of the most abhorrent transgressions of principles we consider fundamental and self-evident. Those involved in editorial activities, be it in our journal publications or conferences, will agree that in recent years we have seen a clear increase in instances of plagiarism, as well as other related ethical violations, perhaps less serious but just as troubling and disruptive.

It is entirely possible that this observation is a consequence of having better means to detect such violations and that their actual frequency was always just as high. The fact, however, is that we are now confronted with this problem. Just as other scientific communities are doing, the IEEE Control Systems Society (CSS) has taken serious steps to address the problem. I would like to use this message to discuss this matter further, differentiate between plagiarism and the "other ethical violations" I mentioned, and to inform our membership about the concrete steps we have taken to deal with them.

For starters, it is instructive to define "plagiarism." A legal definition found in the *Merriam-Webster's Dictionary of Law* is "to copy and pass off (the expression of ideas or words of another) as one's own; to use (another's work) without crediting the source." The etymology of the word is attributed to the Latin *plagiarus*, meaning "kidnapper," but digging a little deeper the root is most probably the Greek word *plagios*, which describes someone acting indirectly or sideways and, by extension, "devious." Although its basic meaning is clear, the way plagiarism is viewed and interpreted varies quite widely by discipline: plagiarism in literature, in journalism, in entertainment, in schools, and of course in academia.

A very interesting article on plagiarism by Clarke [1] (brought to my attention by Dennis Bernstein) explores the many facets of the concept and even provides arguments that have been given in favor of the practical value of plagiarizing. For example, "it is simply impractical to avoid repetition, uneconomic for every author to deliver originality in every element of everything he or she writes, and a waste of time and energy that could be applied to more constructive activities" and "in addition to the impact on authors, readers are poorly served by an excessive emphasis on the avoidance of plagiarism. Citations clutter text. Long reference lists take up space." Although some of these arguments are thought provoking and serve to highlight the complexity of any ethical principle, I will assume that, for our purposes, plagiarism is indeed morally repugnant and unacceptable in a scientific research community.

Although the cases of discernible plagiarism have indeed increased, in my experience they are still relatively few compared to the "other ethical violations" I mentioned earlier. By far the most common of these violations is that of multiple submitted papers by one author (or group of authors) with substantially overlapping material. Note the wordiness necessary to describe this behavior in contrast to the convenient single word "plagiarism." Some have used the term "self-plagiarism" to describe it. Many are opposed to such a term, partly because it is inconsistent with the very definition of "plagiarism," which implies stealing something from "another," not yourself. I

Christos Cassandras exploring the Galapagos sea life.

would argue that *real* plagiarism is so much worse that the term should be reserved just for what it really means.

What exactly do we mean by "multiple submitted papers by one author (or group of authors) with substantially overlapping material?" There are actually many variants of this phenomenon. Sometimes one sees two, three, or more papers by one author (or group of authors) submitted to the same conference with slightly different titles but with as much as 50–80% of the material in each paper being virtually identical. Other times, one accidentally discovers that a paper submitted by an author to a journal is very closely related to a paper submitted to a different journal, with slightly different titles and very minor differences, and that both papers were submitted by the same author at roughly the same time. In the first case, one can argue that the author is simply naïve and ignorant: he/she sees nothing wrong with what they did. In the second case, one can clearly see an element of malice, where the author probably tries to maximize his/her chances of adding a publication to his/her record by trying two different venues (and presumably hoping that there will not be a common reviewer for the two papers).

Regardless of the motive or thought process behind these actions, the end result is that editors and reviewers waste time in processing duplicate work. At a time where our community is heavily burdened with reviewing duties, this waste is by itself detrimental to the mission of any journal or conference program committee. From the standpoint of a reader or conference attendee, being exposed to roughly the same research outcome twice or more is just as wasteful and, to some extent, insulting. What if everyone were to break down a particular piece of research and its results into so many heavily overlapping small parts that it would force all of us to examine each and every part to ultimately extract what is useful?

I will not dwell on some of the other ethical violations sometimes observed.

Let me just mention some examples. Occasionally, one discovers that a coauthor in a paper is actually unaware of the fact that he/she is listed as such. This usually happens when a student or junior researcher includes the name of a well-respected supervising professor or colleague in the hope of receiving more favorable reviews. Another situation arises when authors of an accepted paper (usually for a conference) introduce additional material in their final version, thus bypassing the reviewing process. Since it is extremely difficult to check every final version of every paper, such additions often go undetected, yet giving a reader the impression that they have gone through the scrutiny of peer reviewing.

What happens when someone is "caught" plagiarizing or committing one of the other lesser ethical violations? There seem to be two main lines of defense. The first is claiming pure ignorance: the author simply states that he/she could see nothing wrong, for example, with sending the same paper (or two very similar papers) to two different journals. We have seen this reaction mostly among authors who are relatively young or from countries that have not traditionally been part of the western-style tradition of conducting research. The second is finger-pointing: here, an author blames a coauthor as responsible for the sin committed behind his/her back. The most remarkable and common situation is where a (usually senior) professor points the finger to a student who failed, for example, to properly cite someone's work or submitted two papers with roughly the same material in both. This behavior is doubly reprehensible: not only is an ethical violation being committed, but this person is also willing to throw his/her own student under the bus, refusing to take any ownership of a written piece of research clearly bearing this person's signature in the form of coauthorship. In one recent such instance, a (senior) professor defended himself against a plagiarism allegation by stating that he was too busy with more important

matters to actually read the paper he had coauthored with a student, thus placing the entire blame on the student! This cynicism is stunning, but it also indicates how lightly some people seem to take such matters.

Over the past few years, the CSS has been using software that scans every submitted paper (to one of our conferences, for example) and looks for similarities with documents available in an enormous database of published work, including papers submitted to the same venue. This software is extremely effective at identifying strings of characters shared by two or more documents. Unfortunately, it is a bit too effective: it tags, for example, two papers that share common introductory material or equations that are so standard as to appear in just about everyone's paper in a particular technical area. It is, therefore, dangerous for its high rate of false alarms and demands careful human intervention to filter and properly interpret the outcome of the automated part of this process. As a Society, we have not only put into place mechanisms for such filtering, but we have also created a committee responsible for examining any case of plagiarism or other ethical violation in publishing. The charge to this committee (officially in place in 2012) is not only to identify offenders but, equally important, to ensure that every author is protected against false allegations or poor results from the automated process.

Ultimately, the CSS, along with several other IEEE Societies and non-IEEE scholarly organizations, is determined to protect the intellectual property of its members. At the same time, we are determined to protect ourselves against the waste of valuable resources in reviewing and in clogging databases with duplicate elements of knowledge that end up wasting everyone's time.

Christos G. Cassandras

REFERENCE

[1] R. Clarke, "Plagiarism by academics: More complex than it seems," *J. Assoc. Inform. Syst.*, vol. 7, no. 2, pp. 91–121, 2006.