Narrowing the Readability Gap Between Scientific Papers and the World Wide Web

Michele Catasta
EPFL, Switzerland
michele.catasta@epfl.ch

ABSTRACT

As of today, publications are treated as self-contained entities, with usually a few tens of references to relevant papers in the field. References have a restricted semantic: they can only point to papers as a whole, rather than to a specific portion of the document (as anchor hyperlinks can do with HTML pages). The restriction is due in part to LaTeX–i.e., papers indeed are not hypertexts–although an extension could be developed. Regardless, the convergence between references and Web links can improve the readability of publications as a whole.

For instance, thousands of papers cite MapReduce [1], and most of them re-explain the programming model, rather than pointing the reader to the appropriate section [1:2] of the original paper. Such practice would permit the reader to jump quickly to the intended excerpt, rather than having to scan most of the content.

At the same time, it could also foster conciseness, because the community can consider (by convention) "anchor references" as pointers to must-read resources. Authors will then avoid rewording entire paragraphs, using instead anchor references to provide the needed background information.

BODY

Paper references as anchor hyperlinks can improve the readability and, ultimately, increase the entropy of the scientific corpora.

REFERENCES

[1] J. Dean and S. Ghemawat. Mapreduce: simplified data processing on large clusters. In *Proceedings of the 6th conference on Symposium on Opearting Systems Design & Implementation - Volume 6*, OSDI'04, pages 10–10, Berkeley, CA, USA, 2004. USENIX Association.