

KEEPING THE BALL ROLLING: TEACHING STRATEGIES USING WIKIPEDIA

An Argument in Favor of its Use in Computer Science Courses

Carlos Cotta

ETSI Informática (3.2.49), Universidad de Málaga, Campus de Teatinos, 29071 Málaga, Spain
ccottap@lcc.uma.es

Keywords: Web-based strategies, Wikipedia, Collaborative Work, Sophomore Programming, Computer Science.

Abstract: Edition of Wikipedia articles has been recently proposed as a learning assignment. I argue it ideally suits Computer Science courses, due to the intrinsic mathematical nature of the concepts and structures considered in this field. It also provides benefits in terms of autonomous research and team-working, as well as a valuable legacy for future years' students. This view is supported by a two-year experience in sophomore programming subjects in the University of Málaga, Spain.

1 INTRODUCTION

It is out of question that the younger generations of students perceive issues such as socialization, networking, and access to information in a fundamentally different way than older generations, cf. (Rosen, 2004; Zook, 2007). They routinely utilize the web not just as an integral tool in their quest for information on any topic, but as a form of self-expression: as (Dye, 2007) points out, digital creation is the standard outlet for their creative urges. They are used to, and lively seek, playing an active role in content management, evolution, or annotation – e.g., (Tucillo, 2007). This means that in addition to creating content from scratch, other forms of interaction such as commenting, tagging, etc. are widespread. This behavioral pattern is very well suited to collaborative initiatives such as social news websites, wikis, etc. (Bundy et al., 2007). Regarding the latter, they offer a great opportunity to canalize this content-production pulsion, fostering as well other valuable educational skills, such as team-working among others.

I specifically consider here the case of Wikipedia. Born in early 2001 as a small collaborative project, Wikipedia has become a huge multilingual, collaborative encyclopedia whose contents are often the prime source when it comes to search for basic information on certain topics. Indeed, several studies (Giles, 2005; Rosenzweig, 2006) have shown that the factual ac-

curacy of Wikipedia article is comparable to reputed commercial encyclopedias.

Quite interestingly, scientific areas in general, and computer science (CS) in particular, are typically very reliable in the Wikipedia (Read, 2006) (such fields usually rest on mathematical foundations, and are less prone to biased editing). This firstly implies that Wikipedia constitutes a valuable resource to CS students looking for complementary information. It can be secondly inferred that a well-informed group of editors and administrators exists for this field. This turns out to be important as a quality filter for a teaching strategy that uses the edition of Wikipedia articles as a student assignment. Such a strategy was semi-nally proposed by M. Groom (Brockhaus and Groom, 2007) in the context of environmental courses.

This paper argues in favor of this strategy in CS courses, which constitute an ideal context for its development (due to both the nature of the topics treated and the skills of the students). This is backed up by a two year experience in sophomore programming courses in the University of Málaga (UMA), Spain. As subsidiary benefits of this strategy one can cite autonomous research, team-working, and critical sense (benefits arising from the use of the tool). The primary benefits are the knowledge acquired by students on the topics treated and the valuable legacy for future years' students (benefits arising from the contents developed).

2 ACADEMIC CONTEXT

An experience has been carried out in two subjects taught in sophomore CS courses in the UMA. These subjects are Abstract Data Types (ADT) and Analysis and Design of Algorithms (ADA), two core subjects in the three CS degrees imparted in the UMA:

- ADT is a subject that focuses on the characterization and formal specification of abstract data types. The goal is having the students acquiring an adequate knowledge of the properties and specification of most usual ADTs (stack, queues, trees, etc.) as well as being capable of specifying new ADTs, either *de novo* or by extending pre-existing ADTs. Formal specification is done using Maude, an interpretable language for algebraic specification based on equational logic.
- ADA focuses on the study of techniques for algorithmic design (e.g., divide and conquer, greedy methods, dynamic programming, etc.), providing the students the means for reasoning about their applicability and suitability for specific problems, as well as for actually applying them. The student should be also capable of analyzing the so-constructed algorithms in terms of their computational complexity in space and time, as well as ascertaining the intrinsic complexity of solving specific problems.

As it can be seen, both subjects deal with topics of a strong mathematical foundation, thus naturally adapting to a neutral approach based on mathematical structures and rigorous constructs.

It must be also mentioned that Spanish universities –and the UMA is no exception– are now involved in the process of adapting to the European Space for Higher Education following the guidelines of the Bologna Declaration¹. While the new structure of CS degrees has been delineated but not yet applied (it will be in successive years starting in the 2010-11 academic year), new teaching methodologies adapted to the philosophy of the European Credit Transfer System (ECTS) –in which the satisfactory completion of a subject amounts to successfully acquiring certain skills and competencies– are already being put in practice (Cotta, 2010). These teaching strategies are student-centric and try to boost autonomous (individual and in group) work by the students. The competence-based structure of subjects lends itself very well to a continuous assessment strategy. Thus, the grade is obtained by the outcome of numerous activities during the semester, including the Wikipedia assignment.

¹<http://www.ond.vlaanderen.be/hogeronderwijs/bologna/>

3 ORGANIZATION AND METHODS

The experience has been carried out during academic courses 2007-08 and 2008-09. It has been approached as a team-working activity, and hence students were firstly arranged in groups of two up to four students (the size of the group was later taken into account when assigning topics for edition). The particular composition of groups was a matter of self-organization, since the students grouped according to their own personal affinity and interests.

After groups were organized, a bidding phase was arranged. I selected a list of topics for which Wikipedia entries in Spanish were non-existent or very incomplete. Having focused on the Spanish Wikipedia is due to practical considerations: in general Spanish sophomore students are not fluent enough in English so as to embark on full-scale editing of the English Wikipedia; forcing them to do so would have place an unnecessary burden on them. Also related to this, the legacy of the experience is going to be much more useful for future students who will find the information in their native tongue. Note also that the topics chosen for edition were intimately related to the contents of the ADT and ADA. Table 1 outlines the articles that were considered.

The assignment of topics to groups was done on the basis of each group's preferences, prioritizing smaller groups on some topics, and breaking ties using a FIFO policy. The due date for completing the assignment was the end of the corresponding semester, and students were given some guidelines on how to approach the task:

- A goal was having the Spanish entry of the corresponding article as complete as the English version.
- Bibliographical research on textbooks and Internet was encouraged. A warning was also given with respect to the use of copyrighted material.
- Articles did not need being edited at once, but it was necessary to check carefully (both presentation and correctness) the new material before publishing it.
- Related to the previous issue, spelling, punctuation and writing style had to be carefully revised.
- Crosslinking other articles was essential. In some cases linked articles would not exist in the Spanish Wikipedia. Depending on the case, a stub could be created, or even a reasonable edition of the related article could be attempted.

No specific indications had to be given regarding the actual use of the editor front-end of the Wikipedia,

Table 1: Some topics considered in the experience.

ADT	ADA
2-3 and 2-3-4 trees	Amortized analysis
AVL trees	Backtracking
B trees	Bellman-Ford algorithm
Binary trees	Boruvka algorithm
Binary search tree	Branch and bound
Fibonacci heaps	Cryptography and complexity
Graphs	Divide and Conquer
Heaps	Dynamic Programming
Linked lists	Floyd-Warshall algorithm
Maude	Johnson algorithm
Queues	Karatsuba algorithm
Red-Black trees	P vs NP
Sets and bags	PSPACE
Splay trees	Recurrence relations
Stacks	Shortest path problem
	Strassen algorithm

and the students had no problem when researching their way through the editing procedure.

4 Results

The outcome of the experience can be regarded as very positive. The Spanish version of the Wikipedia is now very complete on the topics mentioned, constituting a very valuable resource for Spanish-speaking students all over the world. No major problems were observed during the realization of the activity with regard to the administrators. In several cases, edited articles departed somewhat from the standard encyclopedic style, being excessively technical and/or direct. In these cases the corresponding entries were tagged as in need of “wikification” according to style conventions.

As mentioned before, lack of fluency in English was one of the reasons why we focused on the Spanish Wikipedia. However, while written expression in acceptable English is somewhat difficult for an average student, their reading skills have been shown in general good enough to be able to adopt material from the English Wikipedia or from other sources. Note also that with some exceptions in which Spanish entries were non-existent, in many cases there was a base article (even if very incomplete) to start working with. This means that a straight translation of a non-Spanish article was not an option in general. On the contrary, a more selective approach to detect missing information or relevant gaps in the Spanish article was necessary. Students were somewhat framed by the context of ADT and ADA courses as well, implying that they

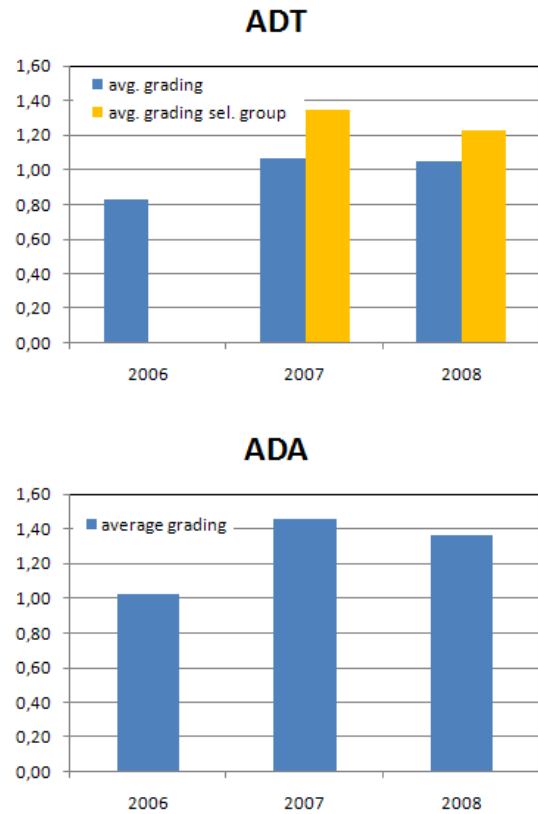


Figure 1: Academic results in the last years.

also added some new material of their own, regarding specific topics (e.g., some Maude specification for certain ADTs).

Figure 1 shows academic results during the last three years (2006 is being used as control). In the case of ADT, the test group (comprising +60 to +70 students depending on the year) is one out of five groups (comprising a total number of between +200 and +300 students) belonging to technical engineering (3-year) degrees. These five groups are course-wise coordinated, down to sharing the same final exam, and hence a comparison is made between the average of these five groups and the test group. In the case of ADA, the test group is the unique group (comprising +30 to +60 students) in the engineering (5-year) degree, and therefore the comparison is done with respect to the same group in 2006. Average marks are computed according to 5 different grades (0=fail, 1=pass, 2=good, 3=superior, 4=outstanding). Notice that ECTS teaching strategies were implanted in 2007-08, and it seems to have resulted in an overall improvement of grades (even if just slightly in some cases). The group that followed this experience obtained remarkably better average gradings. A point of caution is nevertheless required since it cannot be

Table 2: Confusion matrix relating completion of the Wikipedia assignment to successfully passing the course.

		Wikipedia		subtotal
		yes	no	
pass	fail	11.6%	8.7%	20.3%
	pass	69.6%	10.1%	79.7%
subtotal		81.2%	18.8%	

		Wikipedia		subtotal
		yes	no	
pass	fail	21.1%	11.3%	32.4%
	pass	57.7%	9.9%	67.6%
subtotal		78.9%	21.1%	

ruled out that other factors (e.g., different teachers provide different approaches to certain topics, motivation can differ, etc.) take part in the observed distribution of grades.

To obtain a better perspective of the results, Table 2 shows the cross-distribution of passing percentages and completion of the Wikipedia assignment. Following this distribution, we compute the probability of passing the course conditioned to having completed the Wikipedia assignment

$$P(\text{pass}|\text{Wiki}) = \frac{P(\text{pass} \cap \text{Wiki})}{P(\text{Wiki})}.$$

Similarly, we compute $P(\text{pass}|\neg\text{Wiki})$, the probability of passing conditioned to not having completed the assignment. We obtain that among students doing the assignment, 85.7% pass ADT and 73.2% pass ADA. Conversely, among those who do not complete the assignment these percentages are 53.8% and 46.7%, much lower values.

5 CONCLUSIONS

Wikipedia is a prime example of a content provider based on crowdsourcing. It is also one of the most popular references from which basic information is sought at an undergraduate level. Actively using this resource not just as an information source but as a vehicle for canalizing academic content creation is thus a form of positive feedback. It constitutes a way of keeping the ball rolling, contributing to the global information pool. More importantly as to which education concerns, it is an excellent way of promoting self-study of selected topics. The need of com-

municating academic information in a sensible way (and in an self-organized environment with its own control and self-repairing mechanisms) stimulates the student to comprehend the topic at hand. This is particularly well-suited to subjects in which these topics have a mathematical foundation, or at least rest on rigorous facts and/or procedures, since they rely more on factual accuracy and understanding on the underlying math, than on contextual interpretation or personal bias of the student.

ACKNOWLEDGEMENTS

The author is supported by MICINN under project NEMESIS (TIN2008-05941). Thanks are due to the reviewers for useful comments.

REFERENCES

- Brockhaus, A. and Groom, M. (2007). Using Wikipedia to reenvision the term paper. In *EDUCAUSE2007*. (ID: EDU07287) – <http://www.educause.edu/Resources/UsingWikipediatoReenvisiontheT/162173>.
- Bundy, C. B., Cramer, R. J., and Worsham, D. M. (2007). Blogs, wikis, and podcasts: Personal authoring technologies enrich communication and expression for you and your students. In *EDUCAUSE2007*. (ID: EDU07167) – <http://www.educause.edu/Resources/BlogsWikisandPodcastsPersonalA/162292>.
- Cotta, C. (2010). An ECTS experience in sophomore programming courses in the University of Málaga. In *INTED2010 – International Technology, Education and Development Conference*, Valencia, Spain. IATED. (Forthcoming).
- Dye, J. (2007). Meet Generation C: Creatively connecting through content. *Information Age*, May 1.
- Giles, J. (2005). Special report Internet encyclopaedias go head to head. *Nature*, 438:900–901.
- Read, B. (2006). Can Wikipedia ever make the grade? *The Chronicle of Higher Education*, 53(10):A31.
- Rosen, L. D. (2004). Understanding the technological generation gap. *The National Psychologist*, 13(2). March-April.
- Rosenzweig, R. (2006). Can History be open source? Wikipedia and the future of the past. *The Journal of American History*, 93(1):117–146.
- Tuccillo, D. (2007). Standing room only: want to get teens excited about the library? Just surrender some control. *School Library Journal*, 53(3):46–49.
- Zook, G. (2007). Technology and the generation gap. *LLRX.com*, August 27. <http://www.llrx.com/features/generationgap.htm>.