WikiSym 2010 Doctoral Symposium Application

Joseph Corneli

April 6, 2010

1 Introduction

The study of *hypertext* begins in 1965 in and around the Xanadu project ([7], cf. [8]). But where have we gotten to over the years?

At the backend, now we have the relational database model [3] and various triple-stores to be had for free[1], not to mention the rest of the semantic web; and these things can, of course, talk to each other (cf. [2]). In the frontend, we see things like tag clouds, social networking sites, AJAX, and all the familiar aspects of a technology-enhanced daily life.

The education sector has launched itself into this space with tools like Moodle, Elgg, and more recently, with coordinated research on *personalization in technology-enhanced learning*.[2]

As for me, I got interested in this stuff in 2002, when I began to ask *what support can computers offer to people studying mathematics?* It seemed to me that we needed a platfrom like GNU/Linux, but for math, not software. Coincidentally, around the same time I discovered the online community at PlanetMath.org, which shared many of the ideals of the free software movement, and seemed to me a reasonable vehicle to use in pursuit of this overall vision.[3]

As time went by, I realized that many of the tools we needed to make mathematics education work well on computers were really general-purpose tools, combining semantic web ideas with the free software ethos and the online community as a general-purpose learning platform. After doing quite a bit of solo prototyping work[4] I am now seeking to establish myself as a researcher in this field.

2 Central Topic/Question

Can we turn PlanetMath into a place where people can get a good education in math for free?

Let's look briefly at the implications of this question.

[1] E.g. `http://4store.org/`

[2] `http://www.role-project.eu/Deliverables`

[3] `http://planetmath.org`

[4] `http://metameso.org/files/unstable-arxana.pdf`

First, there are some economic implications, as nations are trying to drive the cost of education down, while almost everyone currently bewails the low number of qualified graduates in STEM subjects (science, technology, engineering, and mathematics).

Second, there are logistical implications. Supposing the question was answered in the affirmative: how would learners who had successfully completed their studies on PlanetMath present their credentials to the world? Perhaps a system of examinations like the Tripos, SAT II, Subject GRE, or grad school qualifier exams would become even more important parts of a student's portfolio, with the concomitant issues surrounding learning-to-the-test. Or alternatively, perhaps a "good education" in STEM would bring students closer to the research frontier, and participation in programs like the NSF REU would take the place of in-person courses for many students.

But these points are really moot until we've solved the core problem: assuming we agree it's a good idea, *how are we going to do it?* In essence, this is the topic of my Ph. D.

Before getting into my methodology, I'll mention a third area where this work has implications: it gets at the very nature of computers. Can a computer teach subjects that it does not understand itself on some level? No matter what the answer is, it will be interesting – and I'm excited by the prospect of helping to fulfill part of the grand vision of digital computation that has otherwise been rather slow in coming ([13], [14]).

3 Methodology/Research Approach

Resolving the "how question" calls for a sort of dynamic anthropology, that is to say, a comprehensive approach to understanding the range of issues that apply in learning, developer, and educational communities. Again, I'm assuming we more or less agree about the goal!

In this case, an approach to understanding the "shared context in motion" in which we find ourselves can be found in Nonaka and Toyama's model of knowledge creation [9]. The ontological foundations for this theory come from Nishida (see [1]) and can be usefully compared/contrasted/augmented with Hume (see [5]). In short, these thinkers have provided us with a good way to understand evolving spaces of *noeisis* and *praxis* and their combinations.[5]

This all becomes quite concrete when we apply the Nonaka-Toyama model to understand organizational dynamics. I've begun to do this by considering PlanetMath as a particular candidate for "crowdsourcing education".[6] The Nonaka-Toyama model (presented in Table 1, along with simple glosses from Ken Wilber's "AQAL" theory) can be applied to each dynamic role (e.g. studying, teaching, developing code, partnering with other institutions, seeking certification) to break these down into their component bits. In short, the basic objective in this phase is to create a map of the dynamics that will be involved in addressing the "how" question.

As this develops further, we will should be able to offer a critique of previous efforts at describing mathematical knowledge (e.g. [6]), specifically along the lines of (a) does it really describe the "right" dynamics; and (b) how well does it extend to other related and relevant spaces, like hacking code?

[5] `http://gathatoulie.blogspot.com/2010/03/concerning-direction-of-arrows.html` (my blog).

[6] `http://metameso.org/~joe/docs/crowdsourcing-continued.pdf` (revisions to recent presentation I gave).

Phase	Process	Setting	Gloss
Socialization	tacit-to-tacit	embodied/individual	I
Externalisation	tacit-to-explicit	embodied/collective	We
Combination	explicit-to-explicit	virtual/collective	Its
Internalisation	explicit-to-tacit	virtual/individual	It

Table 1: Nonaka and Toyoma's model of knowledge creation

Everything we learn in these investigations will be fed back in to building tools *and*/*or* community *and*/*or* describing best practices. If we find that students, for example, are unhappy with some aspect of the current system, we should ask "what might be better?", and prepare for answers ranging from

"we need technological change" – which will inform our ongoing work on *tools for tool development*; to

"we need cultural change" – which will inform our understanding of *the social aspects of development.*

In sum, the methodology is based on *mapping dynamics*, inspired by the nascent field of "knowledge cartography" in particular [10], by a body of philosophical literature I've briefly skimmed over above, and by simple and effective tools like Github's Network Graph Visualizer[7] and, of course, wikis and wiki-like systems.

4 Contributions to Theory

I expect I'll be able to contribute something to understanding the "physics of social networks", both via empirical work like that of [12], and by collaborative theoretical work on a discretized mathematical physics. An awesome goal here would be to build a sort of multidimensional version of Zipf's Law, as a way to characterize interactions in spaces with heterogeneous inputs and outputs. The ideas in [4] may suggest one way of entry. (These efforts are framed, at least to some extent, by the separate submission "GravPad", which I hope to present in the demo session at WikiSym 2010.)

My previous work notwithstanding, I have a lot to learn when it comes to hacking. I think it would be great to have a "How to Hack It" book that updates Polya's classic "How to Solve It" [11] with heuristics that work well in a contemporary social, online, setting. Since this book doesn't exist yet, I plan to write it as I go.

Finally, the thrust of all of this is contribute to a better understanding of the "economics of the commons" (more properly *koinomics*[8]); towards a better understanding of the way we imagine "property" (the way we manage resources); and, hopefully, of

[7] `http://github.com/blog/39-say-hello-to-the-network-graph-visualizer`

[8] My father, who is an economist, invented this term to describe "the stuff we often talk about".

the activity patterns that produce lasting happiness.

5 Publications

Existing

- Joseph A. Corneli and Aaron Krowne. A scholia-based document model for commons-based peer production. In Martin Halbert, editor, *Symposium on Free Culture and the Digital Library*, 2005.
- Joseph Corneli. Arxana. A tool for building hackable semantic hypertext platforms. (Work in progress.[9])
- Joseph Corneli. Crowdsourcing Education. Presentation at OU's Open Systems Research Group Seminar Series on Systemic Research and Social Action (Symposium: Social Learning).[10]
- Joseph Corneli. GravPad. A demonstration of a platform for live web annotation and content discovery, submitted to WikiSym 2010.[11]

Planned

- A Survey of "Responsive Open Learning Environments"; to be synced with
- A Survey of Society's Education Supply and Demand.
- A GravPad backend (jointly with Craig Ugoretz). Indeed, ideally this would be developed in detail in a series of mathematics/physics/computation papers with several collaborators, and would deal with clustering, recommendation, and other useful stuff.
- "How to hack it" monograph (represents an important "task dependency").
- Visualization and UI: how exactly should we work with dynamic maps? For example, the idea of viewing a large text corpus as a clickable image, or a stream of miscellaneous data as a song is not new – but there is a lot of room for better integration of these approaches into our routine practice.
- A "PLE IDE". Ideally the PLE and the IDE would fold in to one another, and create a tool that is useful both for learning and for hacking.

6 Feedback/Collaboration Opportunities

I think working with anyone who shares my passion for understanding and improving social systems would be very rewarding. I suspect that at WikiSym 2010 I will meet many other people with views similar to mine. I hope that we will have the

[9] See Note 4.

[10] `http://metameso.org/~joe/docs/crowdsourcing-education.pdf`, and see Note 6.

[11] `http://metameso.org/~joe/docs/gravpad-demo.pdf`

chance to step outside of our own personal "boxes" – and better yet, generate some communication patterns, and even some quasi-institutional arrangements, that help us meet our shared goals! Given where I'm at in my work, it would be both reassuring and highly rewarding to hear someone (or ones) say "I (we) share your views and want to work with you to bring these things into being." Of course, if we aren't working on just the same thing, but can share interfaces, plans, strategies, and the key bits of implementation that allow the things we're working on to fit together well – so much the better!

Bibliography

[1] Masao Abe. Nishida's philosophy of 'Place'. *International Philosophical Quarterly*, 28(4):355–371, 1988.

[2] C. Bizer. D2R map: A database to RDF mapping language. In *Proceedings of the 12th International World Wide Web Conference*, 2003.

[3] C.J. Date and Hugh Darwen. *Databases, types, and the relational model: the third manifesto*. Addison-Wesley, 2006. 3rd edition.

[4] Tim Van de Cruys. A non-negative tensor factorization model for selectional preference induction. In *Proceedings of the EACL 2009 Workshop on GEMS: GEometical Models of Natural Language Semantics*, pages 83–90, 2009.

[5] Gilles Deleuze. *Empiricism and subjectivity: An essay on Hume's theory of human nature*. Columbia University Press, 1991.

[6] Michael Kohlhase. *OMDoc: Towards an Internet Standard for the Administration, Distribution, and Teaching of Mathematical Knowledge*, volume 1930 of *Lecture Notes in Computer Science*. Springer, 2001.

[7] T. H. Nelson. Complex information processing: a file structure for the complex, the changing and the indeterminate. In *ACM '65: Proceedings of the 1965 20th national conference*, pages 84–100, 1965.

[8] Theodore Holm Nelson. *Literary Machines: The report on, and of, Project Xanadu concerning word processing, electronic publishing, hypertext, thinkertoys, tomorrow's intellectual revolution, and certain other topics including knowledge, education and freedom*. Sausalito, California: Mindful Press, 1981.

[9] I. Nonaka and R. Toyama. The knowledge-creating theory revisited: knowledge creation as a synthesizing process. *Knowledge Management Research & Practice*, 1(1):2–10, 2003.

[10] Alexandra Okada, Simon Buckingham Shum, and Tony Sherborne. *Knowledge Cartography: Software Tools and Mapping Techniques*. Springer, 2008.

[11] George Pólya. *How to Solve It*. Princeton University Press, 1945.

[12] Mark Steyvers and Joshua B. Tenenbaumb. The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science*, 29(1):41–78, 2005.

[13] A. M. Turing. Computing machinery and intelligence. *Mind*, 59:433–460, 1950.

[14] A. M. Turing. Intelligent machinery, a heretical theory. *Philosophia Mathematica*, 4(3):256–260, 1996. Posthumous essay reprinted with permission from the Turing estate.