CFP'93 - History-Enriched Digital Objects

by Will Hill and Jim Hollan

Computer Graphics and Interactive Media Research Group
Bell Communications Research
445 South Street Morristown, NJ 07962
(willhill@bellcore.com and hollan@bellcore.com)

Abstract

Recording on digital objects (e.g. reports, forms, contracts, source-code, manual pages, email, spreadsheets, menus) the interaction events that comprise their use makes it possible on future occasions, when the objects are used again, to display graphical abstractions of the accrued histories as parts of the objects themselves. Prototypes we've built to explore this idea (Edit Wear, Read Wear, Email Wear, Source Code Wear, and Vita Service) demonstrate advantages of using objects in the context of their accrued interaction histories. The prototypes also raise complex ownership and privacy issues. In this paper we introduce the idea of history-enriched digital objects, describe a series of prototypes we have implemented, and discuss associated privacy issues.

1. History-Enriched Digital Objects

Recording on digital objects (e.g. reports, forms, contracts, source-code, manual pages, email, spreadsheets, menus) the interaction events that comprise their use makes it possible on future occasions, when the objects are used again, to display graphical abstractions of the accrued histories as parts of the objects themselves. For example, co-authors of a report can see stable and unstable sections (lines of text are marked by recency of changes or amount of editing) and identify who has written what and when. In the case of reading documentation, one can see who else has previously read a particular section of interest. While using a spreadsheet to refine a budget, the count of edits per spreadsheet cell can be mapped onto grayscale to give an impression of which budget numbers have been reworked the most and least. Or in the context of learning unfamiliar menu selections in a new piece of software, the menu itself can depict the distribution statistics of colleagues' previous menu selections in the same or similar contexts1.

The notion of history-enriched digital objects is similar to physical wear. Usage leaves wear. Physical wear is an emergent property and though it generally remains unremarked upon until it causes a problem, it is also tattooed directly on the worn object, appearing exactly where it can make an informative difference.

Consider some serendipitous uses of wear that everyday life presents. The bindings of cheap paperbacks bend and crack in a manner that allows one to find the last page read. The most often consulted pages among many linear feet of an auto parts store catalog are identifiable by smudges, familiar tears, and loose pages. These smudges, tears, and loose pages index information that clerks are most likely to consult. The polished area of an otherwise painted brass door handle shows where others succeeded in grasping it. In kitchen drawers and cabinets, the best recipe cards are often dogged-eared and stained. These examples remind us that wear sometimes encodes useful information. Even a rediscovery of the law of first significant digit distribution due to smudge patterns on logarithm tables has been reported2.

Physics determines the nature of wear on objects in the world. Fortunately, the history of our interaction with certain physical objects sometimes informs and supports us in performing tasks and its worth noting that, unlike most information and support, such history adds little cost. In the case of digital objects not only can we record a much richer history of interaction but also construct a physics of appearance and behavior for those objects purposely designed to inform and support particular tasks. Thus by enriching digital objects with the history of our interactions with them we can do much more than mimic physical wear.

We have built a series of prototypes to explore history-enriched digital objects. We are interested in how such histories might enrich computer-based activities such as reading, writing, and programming as well as how they might improve communication via email and bulletin boards. In addition, we are exploring visualizations of a variety of activity histories such as those captured by source-code management systems commonly used by large programming projects.

2. History-Enriched Digital Object Prototypes

There are many existing computational devices that hint at the prospect of history-enriched digital objects. Automatic change-bars, citation indices, and download counts on computer bulletin boards are examples. In fact, for some thirteen years, members of our lab have been able to request AP News articles by specifying a minimum number of previous readers and thus easily retrieve articles that colleagues have chosen to read.

Most of the prototypes we have developed are discussed elsewhere1. They include:

- **Edit Wear.** Existing editors have been modified to count edits on a per-line basis within user-defined categories such as author or edit session. One way we make that data available is as marks mapped onto scroll bars. One can readily see where edits were performed, who performed them, their age, and amount of time spent on them. Attribute-mapped scroll bars have the very nice property of collocating information display with navigation control points.
- **Read Wear.** We also record the amount of time each line is viewable on the display and support search by user-defined categories permanently associated with read-only documents. For example, one can see who has read various parts of a document. This can be used to identify people who have recently consulted a manual page for a topic of interest.
- **Email Wear.** We record for each incoming message the message header, the time the message is viewable, number of times it is read, its length, time first read, most-recent-time read, as well as the operations performed on messages such as replying, forwarding, saving, and deleting. Groups can filter mail based upon their members average time on message.
- **Source Code Wear.** We use the time-stamped records generated from standard source-code management systems to implement visualizations of the history of additions, deletions, and

modifications. Users can, among other things, easily identify the rapidly changing sections of code (sometimes called churn).

- **Vita Service.** This provides an extensive programmable interaction history mechanism that captures time-stamped actions and current context. It computes and displays summaries such as total time spent on various information items, action frequencies by item-type, and pair-wise item-to-item movement frequencies.

Our prototypes modify document-processing as we usually think about it in two significant ways. First and most importantly, they move reading and editing from the realm of private to semi-public activity. If this were to occur on a large scale, it would represent a significant cultural shift. Second, the use-patterns our prototypes display may occasion conversations that would otherwise not occur because their origin (the use patterns) would not exist.

One interesting question our prototypes raise is why haven't history-enrichment techniques been more wide-spread up to this time. We see a number of technical reasons. First, the techniques multiply storage requirements. To keep a fully categorized fine-grained edit-wear history for a document for just a few authors can require 20 to 100 times the normal storage. The continuing falling cost of digital storage is making such storage requirements less onerous. Second, most often the objects of interest have no identity as objects in a world of flat ASCII files. So assembling, attaching, storing, and restoring object histories represents a large programming effort. The coming generation of permanent object-oriented databases should simplify the creation of history-enrichment services. Third, only recently have human-computer interface designers begun to investigate group-work issues and, so far, that research has emphasized the programmatic control of group processes rather than how shared artifacts might communicate peripherally.

One is starting to see commercialization of history-enrichment techniques in the personal computer market. For example, one can purchase software that will record and report the amount of time spent in each window a user creates3. While there are obvious time-management uses for such data, we conjecture that the real advantages of history-enriched digital objects will come from sharing information with others. Until the range of potential applications of interaction histories have been identified and explored it is going to be difficult to understand the tradeoffs between the advantages that might accrue and the costs that may be associated with the not-so-subtle cultural changes that are involved.

3. Privacy Issues

History-enriched digital objects, like other techniques that record behavior, raise fundamental privacy issues. The ability to maintain very fine-grained histories of our interactions with computer-based applications, and the potential access to that data by others make it exceedingly important to be aware of privacy implications of the techniques we are investigating. We focus on two issues here: who owns interaction history data and how might those histories be shared.

To make the first issue concrete, consider our use of the prototypes described in the previous section. For large portions of our daily activity at our workstations, we and other members of our research group collect data about every document edit, every reading episode, every menu choice, every command, every switch of context, every movement through text, and every compilation of source-code. But to whom and for what purpose are these data available? They exist. Who actually owns them? While our research interest is in exploring the potential benefits of history-enriched objects, we know we must also confront the privacy issues raised by collecting the detailed data we do in our prototype applications.

Based on our experiences, we have taken the position that interaction history data belongs to the user. Individuals should own and control all of their interaction history records. Moreover, individuals should decide what interaction events to record and those decisions are private as well. How might this be possible? One step in accomplishing this is that our software encrypts all of the data it records. Only by providing their private key can data be used or shared with other users. In this form the issues are not different in kind from maintaining other private data on computer systems. That is not to say that control of computerized records of behavior is a simple issue, as forums such as this make clear.

A second step is that our software leaves it up to the individual user exclusively, as to whether or not to record behaviors explicitly and what kinds of behaviors to record. Particular behaviors can be marked to record or not record. All forms of recording can be toggled on and off with simple commands. Furthermore, when history-recording is toggled off and then on, there exist no data in the history from which the toggling behavior can be inferred. The absence of history for the toggled interval is indistinguishable from inactivity.

Maintaining control of private information is complex. We conjecture that a part of an answer may follow from information workers owning, carrying, and working with their own computers and digital storage. The physical control of data that this arrangement implements ensures a base of privacy and security. Network connections to one's personal machine are the only points at which others can easily monitor activity. We envision a future in which, along with the clothes, wallet, pocketbook, watch, and glasses we bring to work, we will also bring our information tools which are then experienced as "personal" in the same way. We are careful now about access to our wallets and pocketbook contents. We will be careful then about access to our personal information tools based on the same physical control that wallets and pocketbooks afford. When personal ownership of the hardware is not possible, approximations to the control that ownership provides might be obtained by renting portable interface machines or perhaps having it provided by unions or employers.

The second issue, sharing of interaction histories, is not yet as crisp as the first. Data might be shared in a variety of ways that have not yet been fully worked out and will surely evolve based on experience. These range from providing the interaction history data explicitly to a colleague without restrictions, to software packages that require a complete transfer of rights to users' interaction histories as a condition of use. The potential to share data anonymously exists, so that the fact of usage can be made available but not the identity of the users. One might choose the anonymous form of sharing with distant colleagues. The summing of use-patterns over groups can offer ambiguity. A person might use the summed netnews reading interaction history of their department or research group to filter and prioritize collections of articles but not know which particular colleague's usage is involved.

We can imagine an economy developing around use-patterns. It might start with discount prices or free upgrades on software packages - if the purchaser is willing to have anonymous summaries of their use-patterns shipped back to vendors. It might then move on to the availability for purchase of exemplary interaction histories as guides to the use of complicated software or large document sets.

Just as in the first issue, there will be tradeoffs between the benefits afforded by use of history-enriched objects and the potential privacy issues that must be confronted. But given that there may be enormous benefits resulting from sharing of detailed records with close friends, colleagues, and from the ability to aggregate those records over larger collections of individuals, the issues of privacy are even further highlighted.

History-enriched digital objects are feasible and may be advantageous. They offer new opportunities for efficiently informing communities of practice and present new dangers for invading personal privacy. There are reasonable arguments that the techniques may provide enough value to cause their widespread adoption. Personal physical control of encrypted use-patterns in small portable information tools is a base upon which to build tools and conventions that may, in the end, comprise new forms of interpersonal and group relationships.

Acknowledgments

We thank Tim McCandless, Scott Stornetta, David Wroblewski, and the other members of our research group for valuable collaborations and conversations on the topic of this paper.

References

1 Hill, W. C., Hollan, J. D., Wroblewski, D., and McCandless, T. (1992) Edit Wear and Read Wear. In: Proceedings of ACM Conference on Human Factors in Computing Systems, CHI'92. ACM Press, New York City, New York, pp.3-9.

2 Weaver, W. Lady Luck. Double Day, Garden City, NY, 1963

3 Window Watch, ASD Software.

Return to CPSR conferences page.	CPSR Return to the CPSR home page.	Send mail to webmaster.