History Browser Research Proposal

Aras Bilgen, David Bingham
November 15, 2005

Georgia Institute of Technology
CS-6451, Human Centered Computing
Professor Elizabeth Mynatt

Introduction

Wikis are simply pages on the web that anyone can edit. Resonant with the free structure of the web, they have found their use within organizations, small and large groups. A very influential use of the wiki idea is Wikipedia, an encyclopedia that anyone can edit. [7] Although there are other knowledge repositories resembling Wikipedia, it set itself apart by decentralizing the editing task to everyone who visits the page.

The distributed editing effort serves to make more prompt corrections and help in connecting articles. In a regular encyclopedia, this is achieved by “references” and “see also” sections, as well as subject indices. Wikipedia, on the other hand, allows anyone to connect any article to another by categories (a tag system for Wikipedia content), series (collections created thematically by users) and by hyperlinks.

Our inspiration for the project was related to the rich meanings created using the linking features. In a web-based wiki, links in articles connect articles to items, which in turn can connect to other items. Over time, each article gains a richer meaning due to the fact that there are more than one trail that leads to the article. The intersection point of the trails help the reader situate the article across different contexts, thus make more connections regarding the material.

The capability of linking items brings into picture a relation between the anchor and the content. In a wiki notion of linked content, the link and the content are in the same medium. Wiki software uses special markup to create hyperlinks to other hypertext documents. The linking, the link, the traversal of the link, and reaching the linked content happens in the same medium. This is not necessarily the case in real world. An anchor can be in the physical world and the content may reside in the digital world (Semapedia [6]), or both anchor and content can be in the physical world (Vietnam memorial and the connections created by left relics).

These linking capabilities afford many new ways of interaction with our physical world. Overlays are one example of the use of these capabilities. By using links going from physical to virtual worlds, informational overlays can be coupled with physical settings. These overlays create multiple trails of meaning by using real world meanings people attribute to things and augmenting it with the semi structured information on the web.

We believe that the same kind of interwoven structure, which incorporates real world, virtual world and inter-world links, can be pushed forward by exploiting the cited aspects of wiki in a physical world setting. The goal of this project is to use a hypothetical device prototype to explore anchor-content relation-

ship and participation, particularly the evaluation of quality and relevance of entries against spatial participation patterns.

Background

Bringing together real world and virtual artifacts has been explored in many fields. There were artistic approaches that mainly built on the anchor-content relationship to create digital overlays on top of the physical world whereas academic projects were more geared towards understanding the fundamental aspects of adding information to the physical world.

Artistic Approaches

Yellow Arrow: Yellow Arrow is a physical tagging system that uses stickers shaped as yellow arrows. [9] The stickers have codes and a phone number on them. The users can attach information to the physical medium by entering data via an SMS message. The users can then retrieve the message by texting the code on the Yellow Arrow to a the number.

Yellow Arrow is created as an artistic community building project upon the significance of the arrow shape. Denoting "This... Check this out... This counts", a yellow arrow sticker is meant to mark things that mean something to people personally. Urban settings and people living in the urban settings make the most use of the project, according to the project site.

Semapedia: Semacodes are 2D barcodes that can be scanned by special software written for camera phones. [6] When a user scans a semacode, the camera phone opens the embedded web browser to a URL linked with the semacode. Semapedia builds on this capability to link real world objects to Wikipedia entries. [6] Users compose semacodes linked to Wikipedia articles in the Semapedia web page and then print the produced semacodes.

It was made to "bring the best information from the internet to the relevant place in physical space". The target audience is similar to that of Yellow Arrow.

eRuv: eRuv is an interesting application that creates an historical overlay. [1] The 3rd Avenue elevated train line in lower Manhattan was once an eruv - a Jewish term for a line that separates private and public space. The historical line is reconstructed using semacodes. People with semacode-enabled phones can retrieve historical content associated with the location by scanning special markers that lead to web pages describing the historical importance of the points along the train line.

The aim of the project was to inform and create awareness about a historically significant piece in the physical world by linking it to this information online. It is accessible to anyone on the street, however, it is of more significance to the Hasidic Jewish community - the ones that are affected by the eruv itself.

Academic Approaches

Real World Wide Web: RWWW, short for Real World Wide Web is an augmented reality project aimed to a general purpose, spatially located, global information space. [3] The system creates informational overlays in real world by bringing together both locally relevant information, such as sights around, as

well as contextually relevant information, such as to-do lists and contacts. The paper mentions embedded location data into the web pages to make them relevant to the physical location. The positioning challenges mentioned in the paper are not relevant to our study, since our prototype does not require a virtual 3D overlay accuracy.

The system follows the ubiquitous computing ideal of moving the interaction off the desktop, thus could be used by anyone who needs contextual data via an augmented reality interface.

Dual-reality Artifacts: Dual-reality artifacts are tokens that can be printed with information. Its aim is to create a link between virtual world and the physical medium of tokens. [4] An IUG, Inter Universe Gateway, is a tool to read and write disks that can carry information linked to the virtual world. The token can then be given away and be read or written by another IUG.

Although the project is relevant in addressing the dual concept of physical-virtual worlds and possible connections, it is not directly relevant to our notion of linking physical and real world. Dual-reality artifacts deal more with the mediation of data between the two worlds whereas we are more interested in how we can couple the two.

Proposed Study

History Browser

The History Browser is a small, handheld device that can be easily stored and retrieved from a pocket, belt, or purse, similar to a modern cell phone. Some of the key technical and interaction aspects of the device will now be discussed to ensure a proper understanding for this study. Coordinate location awareness through a global positioning system allows the device to determine its location in physical space. The resolution of North American GPS is as close as 3 meters, which is sufficiently accurate for our study. WIFI or one of the various cellular-based wireless data connections will also be integrated for the device to receive and send information.

Input methods could evolve with the maturity of the overall system, but a basic means of navigating linear information, such as a scroll-wheel, and a means to input character-based information, such as a small-form keyboard or twiddler would initially be required. Other optional input methods could include a microphone for audio and a camera for both still and motion video. To receive information from the device, it will require a screen of appropriate resolution to view character-based or optional still and motion video. To communicate necessary alerts, the device will include a small speaker and vibration feature.

While most of the features are controlled through user preference and many other abilities could be given to the device, the set for this study will focus on those that relate to the following:

Record: When a user finds themselves at a particular location and has information to share with the Wiki-World system, the History Browser device will facilitate the creation of new entries. The system will automatically denote the date, time, and coordinate-location of the article entry. The user can input their information such as a title, body-content, and a user-signature.

Discover: A user will be able to discover Wiki-World entries through traversing the physical landscape. The device can provide an audio or haptic alert to the user when they are near an area where a Wiki-entry has been made. A search-radius preference can be set to expand or contract the range in which the History Browser detects entries. If a large radius is being used, the device can guide the user to the location in which the entry was made in order to provide precise context. This study will focus primarily on the discovery feature and how coordinate-proximity affects meaning-making.

Annotate: If an existing entry is found and viewed by a user, they will have the option to annotate. While not of primary interest, there are two recording and annotation tracts which are likely. The first is most like the current digital Wikipedia where users debate, edit, and append the content of an entry to increase the total understanding of everyone. The second is similar to linear or threaded comments on sites like Amazon.com or Slashdot.org. The content of these entries would be much more opinionated and include rich bits of information relating to personal experiences and ideas.

Browse: The namesake feature of the History Browser is the ability trace back in time through entries related to a specific location. If a particularly interesting entry is found, the History Browser could be used to step through the author's other entries creating both spatial and temporal links.

The creation of such a system will require at least an initial cache of entries for users to discover and annotate. These could be created by computationally crawling geographical and wikipedia data to match coordinate locations with content. Additional value through links can be achieved at this initial stage by the automated procedure of adding new categories and links. Some study leaders and volunteers in a limited test area could scout and manually enter information. This could also be done in a way to limit the physical scope of the study to a small town or popular area.

Aspects of Physical Wiki

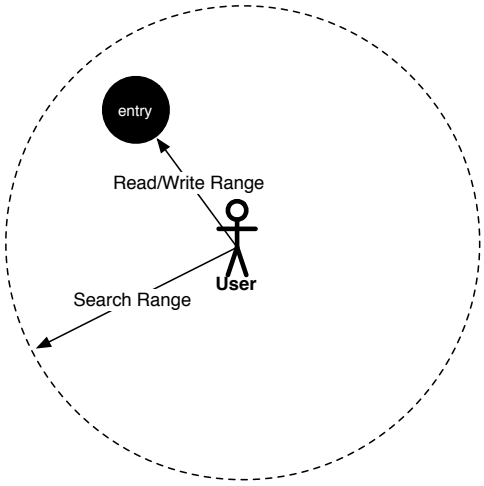
Nicholas Cunningham, the creator of the original wiki at http://www.c2.com put forward twelve principles he sought to satisfy with his software. [8] Most of the aspects on his list were about the editing of the content and nature of the evolution of articles in the system. However, it is interesting to see that we are able to transpose some of the concepts from the virtual world onto the real world. The table below includes the original definitions for the WikiWikiWeb and corresponding definitions.

	WikiWikiWeb definition	Wiki-World definition
Open	Should a page be found to be incomplete or poorly organized, any reader can edit it as they see fit.	Readers can edit poor or incomplete entries and/or add personal insight.
Incremental	Pages can cite other pages, including pages that have not been written yet.	With more citations, historical account about the location will create a complete record.
Organic	The structure and text content of the site is open to editing and evolution.	The community voluntarily explores and contributes to give for rich and robust entries.
Observable	Activity within the site can be watched and reviewed by any other visitor to the site.	Changes in the physical landscape can correlate with the history of entries.

	WikiWikiWeb definition	Wiki-World definition
Trust	This is at the core of wiki. Trust the people, trust the process, enable trust-building.	Credibility will be established for both the authors and individual entries.

History Browser especially focuses on these five aspects. The open, organic form of submission and the observable nature of the entries, both in the physical reality and the virtual domain creates an active consumption-production cycle reinforced by the incremental edits and changes.

In order to understand how users employ these five aspects in various situations, we are suggesting analyze usage and participation patterns using the data captured by feature of the History Browser, proximity alert.



Methodology

Proximity Alert

The proximity alert feature lets a user know when they are within range of a Wiki-World entry. Given the stated hardware set and interaction capabilities, the following raw data points will be logged on the system:

• Search range: (integer, 0 = false)
• Read entry range: (integer, 0 = false)
• Write range: (integer, 0 = false)

As the History Browser alerts the user to an entry within their set search range, the user will have the option of reading and then annotating the entry. When reading the entry, users will also have the ability to rate the content's perceived quality. In each case, the range in which they perform these actions will be tracked for quantitative analysis. By combining the data-log with peer-ratings this study hopes to judge the successfulness of proximity and context in creating quality meaning-making. It is hypothesized that entries with a closer proximity will garner higher peer-ratings creating a more focused Wiki-World entry.

Scenario

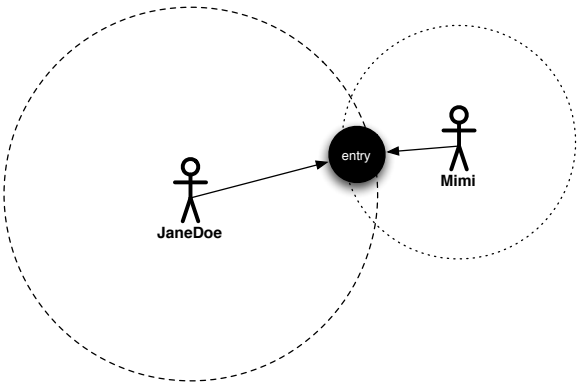
The following scenarios have been written to help understand and put into context the use of the History Browser's proximity alert feature:

Jane is walking near the beach when her History Browser buzzes for her attention. She looks at the display and it informs her of a Wiki-World entry originating 38 meters from her location. This is well within the 50 meter search range preference she set while casually walking in this area, so she checks the entry, which has one annotation:

Fantastic spot to watch a sunset!
Date: August 02, 2005
Time: 19:12

Author: Mimi

It's currently mid-morning, so Jane doesn't bother following the History Browser's directions to the location, but she is very familiar with the area from her daily strolls. She stops, rates Mimi's annotation lower, and jots a quick annotation of her own:



I've lived in the area for 7 years and would recommend heading up the beach to the rocky point to watch a sunset.
Date: November 11, 2005
Time: 10:12
Author: JaneDoe

In analyzing this scenario, Jane has written a convincing annotation and down-graded Mimi's annotation, but she didn't bother to decrease her proximity and put herself in the actual context. Jane's contribution has created 3 pieces of data that can now be analyzed to determine the meaning of this location: The human-created, human-interpreted annotation content, the human-created, machine-logged peer-rating, and the machine-created, machine-logged proximity. While other users could interpret at least these three pieces of information to find meaning, this study will be analyzing the qualitative peer-rating against the quantitative proximity to determine how users attribute meaning to context.

Distributed Cognition

With the collected data, it is expected that the theory of distributed cognition will provide the best framework of understanding. The History Browser, the coordinate location, and all of the ephemeral and permanent artifacts in the vicinity create a rich socio-technical system that is beyond the sole cognition of an individual. As Hutchins comments, "Cognitive science thus concerns itself with the nature of knowledge structures and the processes that operate on them." [2] Similar to Hutchins' own work, the proposed study has created a tool to generate contexts that elicit behavior of users, with a multitude of artifacts, through both time and space.

The most prevalent system in our study is the human-artifact relation. Each user interacts with an an artifact in the real world, reflects on it and maybe writes about it. The instances of interaction are spontaneous, however, both the reflections shared through the system and the changes on the artifact in real world span long time periods. These factors may affect the perceived meaning by users who are reading and writing the entry from different distances as measured by the proximity alert feature. To understand the intricacies and meaning of participation, we need to consider reads and writes over time as a unit of analysis.

Another significant artifact in the system is coordinates. Locations are beyond the pointers to the entries, they are the mediators between the entries and the meaning created in the individuals. Recognizing the entries as artifacts leads to a deeper analysis of the space itself by allowing us to correlate reads and writes, the temporal aspect, with that of the locations.

Study Limitations

A large portion of our study depends on the data logs created by the system. In order to answer the questions outlined in the methodology, we are planning to log locations, different kinds of interactions at each location, editing patterns for multiple users. These data logs can easily become huge due to the granularity of the data. Considering that there is virtually no limit on tagging places and relating them to other content, managing the data logs could be a challenge. Part of this challenge would be storage and data integrity. Another part would be the computational complexity in parsing and displaying data. Neither of these will hinder the outcomes of the study, but should be taken into consideration for experiment designs.

Another possible problem is with the numerous tangential uses of the system. Our anticipated mode of use is read and write entries on everyday events as well as historical data. The infrastructure supporting this task is capable of supporting many other beneficial tasks, such as gaming, focused advertisement services, dating and matchmaking. Furthermore, specialized modes of information browsing for market specific situations, such as restaurant reviews, hotel information, gas prices, real estate information, can be very easily realized using the system. On the malevolent side, the mesh, ad-hoc nature of interaction coupled with the rich location data may benefit terrorists planning attacks. A careful observation on the location data correlated with habits extracted from personal entries can make stalking a problem. Another concern about content is misinformation (due to vandalism or worse, a political agenda) and spamming.

Direct observation is also challenging for interaction with this type of device. Due to the very situated, very spontaneous usage, it may be hard to capture every single aspect of use.

We are acknowledging the existence of privacy concerns such as possible stalking and screening, as well as spamming. The problem of personal privacy can be handled by introducing appropriate measures for anonymity for posting and appropriate measures of authentication for browsing. Users with general web awareness will probably be less impacted with the privacy concerns since there are similar concerns in the world wide web today. entries

Expected outcomes

Intellectual Merit

Previous academic work had addressed issues with combining real world with informational overlays and combining physical objects with virtual content. Several art projects used tags to tie the real world to a virtual one creating overlays of information. Both approaches addressed the contextual nature of the real world content and how virtual content may be enhanced with real world connections that make use of the established context.

History Browser applies the notion of enhancing virtual content by tying it to a relative context to a collaborative body of knowledge where location data is the primary anchor for content. Our work aims to reveal a connection between the proximity that content creation and consumption happens to the relevancy and utility of the content created. Spatial information systems of the future may use the findings from this study as a design guideline to support higher quality content for location contexts.

Broader Impact

The History Browser project explores the possibilities in contextual content consumption and creation. By creating a completely contextual, embedded database of real world items and experiences, the project may create new ways of understanding and participating into meaning making in public spaces.

History Browser as a prototype shaped our study for understanding the anchor-content relations. In the future, further ubiquitous computing research could be accomplished through building a full working prototype for a comprehensive information appliance case study. Such a project will deepen the understanding of modes of situated access, both from a social and a technical point of view. Social studies may include the appropriate uses of the device, the culture built around it or focus on one of the tangential uses of it. Technical analysis may focus on the challenges for geo-location, distributed access, security or data roaming in mesh networks.

References

1. eRuv. http://www.dziga.com/eruv/index.php
2. Hutchins, Edwin, How a Cockpit Remembers Its Speeds. Cognitive Science 19, pgs 265-288 (1995)
3. Kooper, Rob and MacIntyre, Blair. "An Interface for a Continuously Available, General Purpose, Spatialized Information Space."
4. Randall B. Smith, Dual-Reality Objects, The Human Computer Interaction Consortium 2004 Winter Workshop.
5. Rob Kooper and Blair MacIntyre. "Browsing the Real-World Wide Web: Maintaining Awareness of Virtual Information in an AR Information Space."
6. Semapedia & Semacodes, http://www.semapedia.net/
7. Wikipedia, The free encyclopedia. http://en.wikipedia.org/wiki/Wikipedia
8. WikiPrinciples at c2.com. http://c2.com/cgi/wiki?WikiDesignPrinciples
9. Yellow Arrow. http://yellowarrow.net/index2.php

Appendix

Definitions

- Record (v), Annotate, Write: Act of writing a piece of text or attach multimedia content to start an entry, or adding a piece of text or attach multimedia content to an existing entry.
- Record (n), Entry, Annotation: Groups of pieces of text and multimedia content related to a certain location entered into the system by users.
- Explore, Browse, Read: Act of going through the pages in the system. The pages may include ones pointed by the proximity alerts or those linked from other pages.
- Content: Any piece of text, location data, multimedia content and groups of text and multimedia content entered at any time to the system.
- Multimedia content: A general name for images, sounds, videos.
- Structure: The ways in which the pages are linked to other pages.
- Place, Location: A geographical location that can be represented in the GPS coordinate system.

- Person, User: The human that enters information and explores the content on the system.
- Community: Group of users who write, edit and read entries on the system.
- Link: A pointer from an entry or a location to another entry or a location. (This should not be taken as a web page hyperlink by default, there are other hyperlinking possibilities afforded by different hypertext systems)
- Search: To actively look for an entry in a physical sphere limited by the GPS range defined by the user. This is a feature related to the proximity alert.
- Range: The radius of the sphere within which the user wants to receive notifications about the availability of entries.
- Wiki-World: A hypothetical hyperlinking medium for content authoring and retrieval that makes special use of timestamps and location-specific data. Although there could be different methods of access to its body of knowledge, History Browser is the assumed way to access the system for our study.

Aspects of Wiki

- **Open** - Should a page be found to be incomplete or poorly organized, any reader can edit it as they see fit.
- **Incremental** - Pages can cite other pages, including pages that have not been written yet.
- **Organic** - The structure and text content of the site is open to editing and evolution.
- **Mundane** - A small number of (irregular) text conventions will provide access to the most useful page markup.
- **Universal** - The mechanisms of editing and organizing are the same as those of writing so that any writer is automatically an editor and organizer.
- **Overt** - The formatted (and printed) output will suggest the input required to reproduce it.
- **Unified** - Page names will be drawn from a flat space so that no additional context is required to interpret them.
- **Precise** - Pages will be titled with sufficient precision to avoid most name clashes, typically by forming noun phrases.
- **Tolerant** - Interpretable (even if undesirable) behavior is preferred to error messages.
- **Observable** - Activity within the site can be watched and reviewed by any other visitor to the site.
- **Convergent** - Duplication can be discouraged or removed by finding and citing similar or related content.
- **Trust** - This is at the core of wiki. Trust the people, trust the process, enable trust-building.