Generating Educational Tourism Narratives from Wikipedia

Brent Hecht*, Nicole Starosielski, Drew Dara-Abrams*****

* University of California, Santa Barbara – Department of Geography
** University of California, Santa Barbara – Department of Film Studies
*** University of California, Santa Barbara – Department of Psychology

{bhecht, n_star}@umail.ucsb.edu
dara-abrams@psych.ucsb.edu.

Abstract

We present a narrative theory-based approach to data mining that generates cohesive stories from a Wikipedia corpus. This approach is based on a data mining-friendly view of narrative derived from narratology, and uses a prototype mining algorithm that implements this view. Our initial test case and focus is that of field-based educational tour narrative generation, for which we have successfully implemented a proof-of-concept system called Minotour. This system operates on a client-server model, in which the server mines a Wikipedia database dump to generate narratives between any two spatial features that have associated Wikipedia articles. The server then delivers those narratives to mobile device clients.

1. Introduction

Heritage/cultural tourism, ecotourism, agritourism, and other types of information-centric tourism have been of increasing importance in recent years, both to the tourist and to the tourism industry. However, these types of tourism have not yet fully taken advantage of the dramatic increase in available data that has been a hallmark of the Information Age, particularly with respect to the adoption of mobile technologies (Brown and Chalmers 2003). One aspect of tourism that is particularly in need of better mobile device applications is that of educational tourism. Most mobile technologies aimed at the tourism market are either (1) tourist tools can do nothing more than inform users of optimal or nearby tourist facilities or attractions and thus can do little educating (Kim et al. 2004, etc.), or (2) educational devices that are limited in scope due to the amount of required custom content development (Isbister and Doyle 2003, etc.). With Minotour, we attempt to fill this vacuum, and our central methodology is the employment of intelligent narrative technology, particularly data mining techniques informed by narrative theory.

Why use narrative? Much research has concluded that humans have an inherent predilection towards narrative approaches (Mateas and Sengers 2003). In addition, aside from the innateness of narratives to the human experience, narrative has been shown to play a particularly important role within the two constituent fields of our educational tourism test platform - education and tourism - as well as in the combined platform itself. Wells (1986), Mott et. all (1999), and many others have identified the myriad benefits of narrative to education and Gretzel and Fesenmaier (2002) have done the same for tourism. In the educational tourism context, Lanegran (2005) concluded that a high-quality educational tour is one in which a cohesive story is woven while traveling through the landscape. A similar assumption is made in (Isbister and Doyle 2003).

Minotour utilizes Wikipedia as its primary data source, thus eliminating the need for *any* knowledge base development, which, as noted above, is a rarity in the world of automated tour guides with an educational focus. Wikipedia is an online, user-contributed encyclopedia that is available in more than 100 languages and has more than 1.7 million articles in its English version and over 500,000 in its German version, the two languages our implementation supports. Through the use of the Wikipedia corpora, Minotour inherits all of Wikipedia's user-friendly advantages, such as democratized and free information that is global in scope. We also utilize several other unique properties of Wikipedia, as described in section three. However, the text of Minotour's generated narratives – derived directly from Wikipedia – obviously also exhibits Wikipedia's well-publicized disadvantages, most notably with regard to concerns about accuracy, motives, uncertain expertise, volatility, coverage, and sources. (Denning et al. 2005).

In its current implementation, Minotour operates on a simple premise: with a mobile device, the user selects a destination and a narrative tour from the user's current location to that destination is provided. The user's position is determined via GPS and the tour is generated on a server and delivered back to the client. These tours are designed to be experienced *while* the user moves to the destination (in our prototype, this is done via text display; we are also exploring other options, including audio). An example tour with Berlin's Brandenburg Gate as the start feature (feature *a*) and Unter den Linden as the end feature (feature *b*) is provided in figure 1. In this paper, we will walk

Brandenburg Gate

Finally, when the Berlin Wall fell in 1989, the gate symbolized freedom and the unity of the city. It re-opened on 22 December 1989 when the West German Chancellor Helmut Kohl walked through to be greeted by the East German Prime Minister, **Hans Modrow**.

Hans Modrow

After the German reunification, he served as a member of the Bundestag and currently serves as a member of the **European Parliament**.

European Parliament

The European Parliament (formerly European Parliamentary Assembly) is the parliamentary body of the European Union (EU), directly elected by EU citizens once every five years. Together with the Council of Ministers, it composes the legislative branch of the institutions of the Union. It meets in two locations: **Strausbourg** and Brussels

Strausbourg

A revolution in 1332 resulted in a broad-based city government with participation of the guilds, and Strasbourg declared itself a free republic. The Straußburger Münster was completed in 1439, and became the World's Tallest Building, surpassing the Great Pyramid of Giza. During the 1520s the city embraced the religious teachings of **Martin Luther**, whose adherents established a University of Strasbourg in the following century.

Martin Luther

A minority viewpoint disagrees with the attempt to link Luther's work causally to the rise of Nazi anti-Semitism, arguing that it is too simplistic an analysis...Johannes Wallmann, professor of church history at the **Humboldt University of Berlin**, writes that Luther's writings against the Jews were largely ignored in the 18th century and 19th centuries.

Humboldt University of Berlin

Its main building is located in the centre of Berlin at the boulevard **Unter den Linden**. The building was erected by Prince Heinrich of Prussia. Most institutes are located in the centre, around the main building, except the nature science institutes, which are located at Adlershof in the south of Berlin. The University continues to serve the German community.

Unter den Linden

Unter den Linden...is a street in the centre of Berlin, the capital of Germany. It is named for its Tilia or lime trees (also known in North America as basswood trees) which line the grassed pedestrian mall between the two carriageways. Unter den Linden runs east-west from the Brandenburg Gate in the west to the Schlossbrücke (Castle Bridge) over the River Spree in the east. The major north-south street crossing Unter den Linden is Friedrichstrasse.

Figure 1 – *A sample narrative tour generated by our prototype version of Minotour. In this tour, Brandenburg Gate is spatial feature* a*, Unter den Linden is spatial feature* b*, and s = 7 (see section three for details about variables).*

through the theory and methods that are used to generate these tours. The second section provides a quick summary of related work in narrative intelligence. In the third section, we discuss the Wikipedia context of the project and the operation of the narrative generation algorithm. In the fourth section, we outline our understanding of narrative theory and the two key narrative cues, unity and progression, that underlie our conceptualization of the ideal narrative function, the function on which the narrative generation algorithm is based. In the fifth section, we detail the implementation of the project. Finally, we close with concluding thoughts and a description of our plans for future work.

2. Related Work in Narrative Intelligence

In the narrative intelligence framework laid out by Mateas and Sengers (1999) and Mateas and Sengers (2003), the Minotour narrative generation system draws most heavily from the then-state-of-the-art (1999) body of work designed to support human narrative intelligence. Indeed, the key presumption in Minotour is that there is inherent value - increased cohesion, improved recall and comprehension, the ability to overcome certain cognitive obstacles, increased synergy with how tourism is experienced, closer alignment with geography education, etc. - in providing information to users in a form that is easier to interpret as a narrative than in the dominant manner in which Internet information is accessed. In the context in the 2003 version of Mateas and Sengers' framework, this trait of Minotour places it in the Narrative Interfaces category, as Minotour essentially provides a narrative-based interface to Wikipedia, albeit a unique and highly spatial one.

In order to provide narrative support, however, we make a concrete assumption as to the definition of narrative, at least in the context of our system. This definition, outlined in section 4, is used to generate stories, thus placing our work also well within Mateas and Sengers' Storytelling Systems category of work, and within the story-centric sub-category.

While Minotour's feature base lies firmly in narrative support and Storytelling Systems, Mateas and Sengers' Interactive Fiction and Drama category also applies to Minotour. Users interact with Minotour in both spatial and non-spatial ways: they are only allowed to obtain tours beginning in their present spatial location and end destinations are chosen via the medium of the mobile device. In addition, because Wikipedia is editable by any user, a measure of interactivity pervades all aspects of the tours. This feature of Wikipedia, along with our system's ability to bring out emergent features of the link structure and raw text of the encyclopedia, allows any user to influence the narrative that she and all other users experience.

3. The Narrative Generation Algorithm

In order to understand Minotour's narrative generation algorithm - the computational methodology we use to

generate tours such as that in figure 1 - it is necessary to first highlight some properties of Wikipedia.

3.1 The Wikipedia Context

From a geoinformatics data mining perspective, Wikipedia articles can be split into four categories, two of which are critically important here: articles that can be referenced to a spatial entity and articles that cannot. We refer to the former type of articles as "3D articles" because they exist in both the two-dimensional space defined by the spatial reference system used in Wikipedia (World Geodetic System 84) as well as in Wikipedia space, and the latter as "non-3D articles", because they have no spatial reference. Of course, within the context of our educational tourism test bed, the start and end articles of our tours must be 3D articles. 3D articles are identified through user-embedded latitude and longitude information and a basic georeferencing process.

Secondly, it is important to point out that a key benefit of using the Wikipedia corpus is that the online encyclopedia has a large hyperlink structure between articles. We take advantage of this structure in every step of our algorithm. The English Wikipedia had 32.1 million links as of October 2006 (Zachte 2007), with nearly all links representing some kind of meaningful semantic relationship.

Finally, because Wikipedia is collaboratively edited and the average Wikipedia article is contributed to by at least 7 different authors (Buriol et al. 2006), the paragraphs in Wikipedia tend to be more disjoint and contain fewer references to each other than those in other plain text corpora. Wikipedia's encyclopedic writing style also contributes to this phenomenon. As such, we are able consider paragraphs as individual entities called "snippets" and re-arrange them to our hearts content without destroying their semantic value. Because of our current mobile device delivery platform, we have further constrained the definition of valid snippets to be Wikipedia paragraphs between k and l characters in length, with k and l currently set to 200 and 600.

3.2 Operation of the Algorithm

Concisely, our narrative generation algorithm operates with the following goal:

Identify the narrative n_{best} of length s from 3D article a to 3D article b, where $Q(n_{best}) = max(Q(n))$ for all n of length s between a and b and Q is the pre-defined "narrative evaluation function". (n is a path through Wikipedia's link structure between a and b; s is the number of snippets in the desired tour and is derived from the distance between a and b in geographic space.)

Between nearly all pairs of 3D articles for all s greater than a small number, there are dozens to thousands of possible paths through the Wikipedia graph, even when the path length restriction of s is implemented. We identify these paths with the algorithm shown in figure 2. The algorithm is given as input 3D article a, 3D article b, and s.

Conceptually, the algorithm works as follows: Starting from a and b (two separate trees), the algorithm identifies all of the out-links and in-links between a and b and other Wikipedia articles. An out-link is a link pointing from the article being considered, x, to a different article, y; an in-link is a link from y to x. We have determined that these types of links have equal semantic value for our narratives. Once the algorithm has located the destinations of the out-links and origins of the in-links, it examines the links of these destinations and origins. It continues

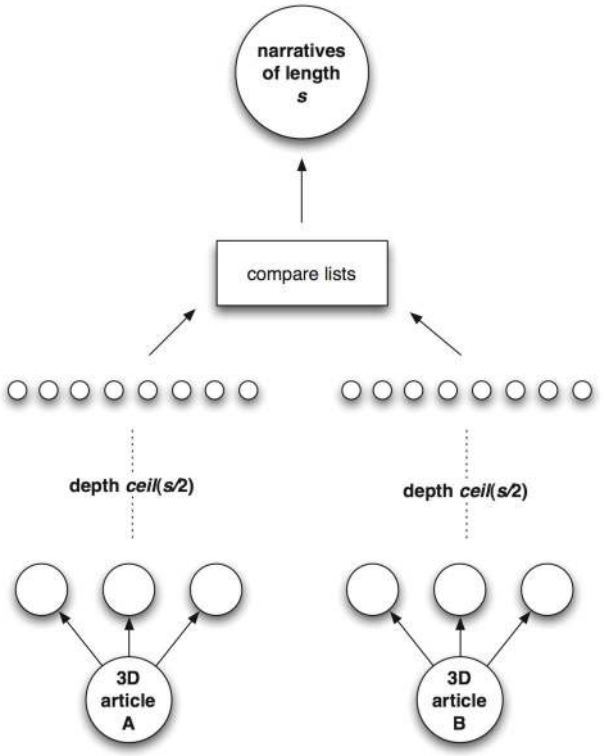

Figure 2 – *A diagram depicting the operation of the path finding portion of the narrative algorithm. Once found, the narratives of length* s *are fed into the* Q *function, and the narrative with the optimal* Q *output is selected and delivered to the user. The above example is for the case of an odd* s. *The algorithm is slightly different for even* s.

recursively until level $ceil(s/2)$ or $ceil(s/2) + 1$ is reached, depending on whether s is even or odd. The lists of articles at the leaves of each tree are then compared. Any article that appears in both lists represents a connection between the trees, and thus a path between a and b. Once identified, the paths between a and b are assessed using the narrative evaluation function Q, and the path with the optimal output is returned. Q(n) is defined as the total error (RMSE) of path n to a predefined ideal narrative function Qi, shown in

figure 3. *Qi* is two-dimensional, with snippet number on the *x*-axis and snippet host article in-links on the *y*-axis. Why have we defined the ideal narrative function in this way? The answer to this critical question lies in our novel use of narrative theory, and is the subject of section 4.

While the implemented algorithm is derived from the conceptual algorithm described above, there is one key difference between the two. If the algorithm were to examine all of the possible links between each article, it would quickly become extremely slow and resource intensive. This problem is well-known in computer science and is related to each tree's "branching factor", or number of new articles to be examined per article. It is easy to see that if the first article has 900 total links (say, 800 in-links and 100 outlinks) and each of those 900 linked articles have 900 of their own links, the amount of articles to be processed becomes untenable ~(900^(*s*/2)). To solve this problem, we have turned to an effective, albeit basic and non-optimal solution that uses a simple heuristic. At each article, we only examine *BF* links, where *BF* is set to an arbitrarily low number that experimentally maximizes effectiveness while also considers the limited computational resources of our prototype environment. We have had success with $3 \leq BF \leq 10$. Taking a hint from A* (Hart et. al 1968), we select the *BF* articles by sorting the in-links and out-links by their likelihood to produce narratives that will result in high narrative evaluation function (*Q*) values. We have found that this approach exhibits results that, while non-optimal, are more than satisfactory.

While we try to keep the narrative dependent entirely on the properties of the Wikipedia link graph, we explicitly prevent one key group of articles from appearing in the narrative: purely temporal articles. This is the third geoinformatics group of Wikipedia articles, and is disallowed from the narratives because purely temporal articles almost always exhibit uninteresting and non-educational semantic connections with their neighbors in Wikipedia space. For example, the article "1979" is essentially a list of events that occurred in 1979, a list that is so disparate that it includes the acquisition of home rule for Greenland and the premiere of "Morning Edition" on the United States' National Public Radio. Other purely temporal articles include "1970s", "April 29", and "11th Millenium".

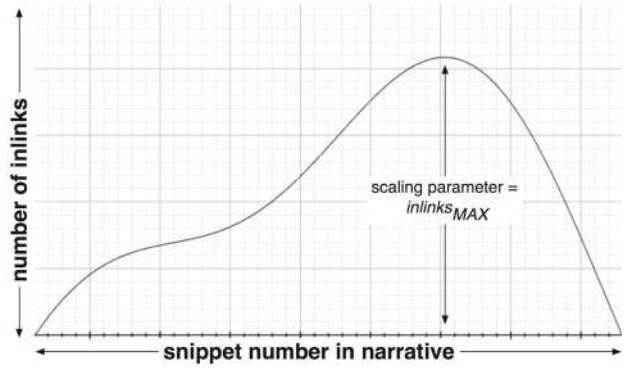

***Figure 3** – the Ideal Narrative Function,* Qi. *The narrative algorithm chooses the path* n *that most resembles this function.*

4. The Idealized Narrative Function

In the previous section, we show that our narrative algorithm's functionality can be simply described as trying to find the path *n* through the Wikipedia graph between *a* and *b* that is most similar to an idealized narrative function, *Qi*. In this section, we discuss the two characteristics of narrative experience – unity and development - that we model in this function, as well as our reasoning for choosing these two characteristics. We hypothesize that when these characteristics are successfully embedded in a generated Wikipedia text, the resulting textual object will be read as a narrative, and the travel from *a* to *b* will be considered a narrative experience.

4.1 Narrative Approach

In their review of narrative intelligence research, Mateas and Sengers (2003) note that, with narrative being such a rich and complex concept, it is critical for narrative intelligence scholars to be explicit about their approach. As an interdisciplinary team bridging the sciences, engineering and the humanities, our approach to narrative is based in narratology, a field of narrative theory developed from the analysis of literature, film and other media. Much work that is adopted from this field draws from structuralist authors and their rigorous interpretation of underlying narrative structures. Structuralist approaches translate these underlying structures into sets of variables, codes, and functions that reside in the text. One of the earliest examples of this approach is Vladimir Propp's *Morphology of the Folktale* (1928), which delineated 8 basic character types and a series of possible functions that each character could enact. Structuralist theories assume that one can define narrative as a text that aligns to a certain set of rules, such as the depiction of characters, plot, and cause-effect relationships. The problem with adopting this model in narrative intelligence research is that these approaches often sidestep the psychological and spatial conditions that influence narrative experience.

We draw our narrative approach, rather, from recent research in narratology that is influenced by cognitive science, including the narrative theories of David Bordwell and Edward Branigan. These theories are grounded in the same rigorous interpretation of structures as Structuralism, but rather than defining "narrative" as the sum of these structures, they locate narrative as a perceptual activity. Edward Branigan writes that narrative is a "perceptual activity that organizes data into a special pattern which represents and explains experience" (Branigan 1992). It is

this understanding of narrative that we use to develop our ideal narrative function. Rather than trying to model a text out of Wikipedia, which possesses a certain set of traits (such as characters who have goals), our narrative algorithm shapes a disparate text corpus according to characteristics of a desired narrative experience.

According to Bordwell, a precondition for understanding a fiction film as narrative is the unity of the text as a formal system. He writes that, in the most formally unified films, "every element present has a specific set of functions, similarities and differences are determinable, the form develops logically, and no element is superfluous. In turn, the film's overall unity gives our experience a sense of completeness and fulfillment" (Bordwell 2006). Thus, for a reader or viewer to understand the narrative as unified, they must perceive the individual elements as interrelating and nothing must seem "out of place." As is also noted in this passage, a sense of development is also key to a sense of unity. Bordwell defines formal development as a progression moving from beginning through middle to end (Bordwell 2006).

Simply including unity and development, also termed "progression," will not designate a given text as a narrative. However, they are two of the most common narrative cues that elicit a narrative mode of organization. Bordwell delineates this narrative comprehension process as follows: cues are organized "to encourage the spectator to execute story construction activities. The film presents cues, patterns and gaps that shape the viewer's application of schemata and the testing of hypothesis" (Bordwell 1987). We hypothesize that incorporating these narrative cues of unity and progression into a representation of Wikipedia will encourage the reader to execute story construction activities and to utilize narrative schema in making sense of the individual snippets.

4.2 Unity in the Ideal Narrative Function

Taken as a whole, Wikipedia is a disparate collection of facts (and some opinions) with no inter-article coherence. However, as noted above, unity is a critical narrative cue. As such, unity must be an important characteristic of the ideal narrative function, the function against which all possible narratives between two 3D articles are evaluated. Of course, the most obvious way that we provide unity is by linking 3D article *a* to 3D article *b* through a series of other articles. Before the user reads the narrative, these articles were likely perceived as distinct entities; afterwards, they are inherently connected. But this form of unity is weak and only informs our ideal narrative function in the most basic way.

More significantly, we provide unity by highlighting the themes in the user's space. This approach to unity is informed by our educational tourism focus. A critical element of geography education is the highlighting of themes embedded in the geography of a region. In fact, a key element of regionalization – a backbone of the United States National Geography Standards (National Geographic Research & Exploration 1994) – is being able to reduce the complexity of a diverse area through the construction of thematic regions.

While it might seem that incorporating a number of themes might disconnect the narrative, the presence of multiple themes can unify the narrative experience as a whole. Each theme serves to draw a conceptual link between two or more snippets. For example, in our sample narrative, issues about German freedom, unification, and community are directly referenced in the snippets about the "Brandenburg Gate," "Hans Modrow," "Strausbourg," and "Humbolt University of Berlin." The themes of German freedom, unity, and community thread together these snippets. On the other hand, the intersection of the academic institution with political and religious forces is present in the snippets from "Strausbourg," "Martin Luther," and "Humbolt University of Berlin."

The need to quantify the thematic requirement is the reason we chose in-links as the defining variable in the ideal narrative function. In the context of Wikipedia, in-links are a highly accurate proxy for generality. When an author working on one entry links to another entry, that author is essentially saying "this entry is important to my entry". Articles that are more important to more entries are more general articles. For instance, in the German Wikipedia, the article for "Poker" has far more in-links than the article for the "1990 World Series of Poker". We make the assumption - proven to be mostly true experimentally - that more general articles have more thematic content. Given the in-links/generality/theme relationship and the ideal narrative function's focus on high in-links articles, the reason for the appearance of themes in our ideal narrative function becomes clear. It is important to note that we include in the narrative function a maximum number of in-links that is less than the actual maximum number of in-links. We found through experiments that articles with extremely high in-link totals are too broad to carry much interesting thematic significance. The most common example of excessive in-links are the articles about days of the year, i.e. the "October 5" article. In the English Wikipedia, this article has thousands of in-links, but has little to no thematic content.

4.3 Development in the Ideal Narrative Function

If we were just focused on theme, it would be ideal to find a set of articles with a high number of in-links (but not too high) and present these to the user. However, it is critical to the narrative foundation of our tours that the reader understands the snippets of the text as developing in a certain direction.

As was the case with unity, there is an obvious answer to the question of how our ideal narrative function incorporates development: each snippet is linked to the next snippet. However, just as with unity, the function also includes conceptually deeper models of development. We accomplish this by incorporating a *small* positive slope in the first two-thirds or so of the function. In this part of the function, each snippet becomes only slightly more general than the next. As such, the reader gains a sense of movement toward generality. She can then question and form hypotheses about the broader themes of her space and how they will connect her start to her destination. This

***Figure 4** – A screenshot from a software emulation of our Minotour handheld client.*

activity mirrors the experience of the traditional narrative reader. In the traditional narrative text, conflict builds to a climax, followed by a resolution. In the context of our ideal narrative function, this conflict is manifested in the question "What does this snippet from this thematic article have to do with the specific space in which I am moving"? As such, before the user is returned from wiki-space to a 3D article, the broadest theme of that space – the climax – is revealed. Then, the movement from the broadest theme back to the concluding 3D article – the resolution – provides a sense of completeness and closure to the experience. In the case of our sample narrative in figure 1, the climax occurs at the Martin Luther snippet. At this point, we expect that the user is maximally wondering what a highly broad article like "Martin Luther" has to do with the specific space in between the Brandenburg Gate and Unter den Linden. This curiosity is quickly satisfied in a mere two hops through the Wikipedia graph, as the user learns of the connection through Humboldt University.

Since Minotour's generated texts are designed to be delivered as the user travels from feature *a* to feature *b*, they are accompanied by the plotline the user is experiencing by moving through the space. The progression of this two-trajectory context in which the stories diverge at the beginning (at *a*) and rejoin at the end (at *b*), is one that the user is used to perceiving within a narrative, and thus aids development. Thousands of examples of this context exist in pop culture media. For instance, nearly all Seinfeld episodes begin with one narrative that splits into two (or more) at the beginning and intersects at the end ("The Limo", etc.)

5. Implementation

We have taken a basic client/server approach to our implementation, an approach that maximizes the individual flexibility of the client and the server. All of the narrative generation work takes place on the server, while all of the narrative delivery is done by the client. When a narrative is desired, the client (currently a Windows Mobile 5 application still in very early stages of development), sends the article id number of 3D articles *a* and *b* to the server, and the server returns the optimal narrative between these two features.

The server operates on information provided by a MySQL database of Wikipedia data parsed by a custom Wikipedia processor designed to operate on the semi-regular Wikipedia dumps provided by the Wikimedia Foundation, the operator of Wikipedia. In addition to extracting links, snippets, and other critical data, this parser is responsible for identifying the existing geo-information in Wikipedia, as well as doing the rudimentary georeferencing and tagging of purely temporal articles. We have written the parser to support any of the hundreds of languages for which a Wikipedia version exists, but so far, only the English Wikipedia and the German Wikipedia (the two with the most articles) have been tested.

The client side of this project is the lesser-developed side, and we are exploring delivery platforms other than handheld devices. That said, our current handheld software, developed in C# and the Windows Mobile Native C++ API, is effectively a mini-geographic information system (GIS) with GPS capabilities, customized for the

various features and desired user experience of Minotour. Because no suitable open-source code could be located, it was necessary to write this software from scratch. A screen shot of a software-emulated version of the application can be found in figure 4.

6. Conclusion and Future Work

In this paper, we have presented a novel, narrative theory-informed approach to data mining from a Wikipedia corpus within the context of educational tour generation. While we have completed a proof-of-concept system, we have many more research avenues to explore.

First and foremost, we intend to rigorously investigate both theoretical and artistic methods for increasing the narrative pleasure of our generated narratives. One idea currently in the works is eliminating spatially-referenced articles (in addition to temporally-referenced articles) from the body of the narratives. Spatially-referenced articles play a critical role as the start and the end points of our narratives. However, we have found that when they have snippets that appear in the body, these snippets suffer from the same semantic weakness problems as purely-temporally referenced articles. For instance, in figure 1, the implicit "lies on" relationship between Humboldt University and Unter den Linden is a rather uninteresting and non-educational one, and one that would be obvious to any user who looked at the map on their mobile device client. Other possible avenues for increased narrative pleasure include providing theme-based tours (by utilizing Wikipedia's category graph) and using Wikipedia's link structure to include more articles in the narratives that are closely tied to the space in which the user is traveling.

Secondly, we are exploring different narrative theoretical constructs with which to examine and improve our generated narratives. We aim to develop further research in the area of cognitive science-based narrative theory with the goal of proposing better strategies for the interdisciplinary adoption of contemporary narratology.

We are also excited about the possibilities for client-side development. Raw text presentation is not the best delivery format for our narratives. We are exploring many possibilities for ways to utilize the presentation as a means to aid narrative interpretation, either through the increase of unity and development, or via other theoretical constructs. One idea on the table includes audio (think "location-based podcasts" [Schöning et al. 2007]). We are also looking into turning the faceless narrator of our stories into a more explicit character, perhaps even similar to that in (Persson et al. 2003) or (Isbister and Doyle 2003). We have many ideas for using the properties of Wikipedia to enact Persson's suggestion to closely integrate character behavior and presented content. This might include placing greater emphasis on the real life Wikipedia users who input content, as well as considering article conflict statistics.

Following (Persson et. al 2003), we are taking steps towards designing an appropriate evaluation for Minotour. We will determine the average user's reliance on narrative schema for the comprehension of a random series of linked snippets (our baseline) as opposed to Minotour generated narratives.

Finally, we also plan on implementing spatial feature contribution functionality. Ideally, users should be able to enter information about a spatial feature on which Wikipedia has no information and, immediately afterwards, receive a tour with that feature as the start or end destination. This would add a whole new level of interactivity to our current system.

Minotour is currently heavily customized towards its educational tour narrative generation test platform. However, the ideas behind the implementation can be utilized in other areas and with different goals. Most apparently, the concept can be applied to generating narratives between *any* two Wikipedia articles, spatial or not. In initial tests, the authors learned a lot from narratives generated between biographies and even between two mathematical concepts. More broadly, we are considering how this idea could be employed on the Internet as a whole. While we have not explored the narratology implications of such non-spatial and/or non-Wikipedia applications, these are vital further research directions.

6. Acknowledgements

Research supported by National Science Foundation (NSF) IGERT Program in Interactive Digital Multimedia (Award #DGE-0221713). Special thanks to Julie Dillemuth, Kirk Goldsberry, John Roberts, Dr. Keith Clarke, Dr. Tobias Höllerer, Dr. Annemarie Schneider, Dr. David Lanegran, Johannes Schöning, and Meri Marsh for their assistance on this project.

7. References

Bednarz, S. W., Bettis, N. C., Boehm, R. G., De Souza, A. R., Downs, R. M., Marran, J. F., Morrill, R.W., Salter, C.L. 1994. *National Geography Standards*. Washington, D.C.: National Geographic Research & Exploration.

Bordwell, D. 1987. *Narration and the Fiction Film*. London: Routledge.

Bordwell, D. and Thompson, K. 2006. *Film Art: An Introduction*. Boston, MA: McGraw Hill.

Branigan, E. 1992. *Narrative Comprehension and Film*. London: Routledge.

Brown, B. and Chalmers, M. 2003. Tourism and mobile technology. In *Proceedings of the Eighth European Conference on Computer Supported Cooperative Work*. Helsinki, Finland: Kluwer Academic Press.

Buriol, L. S., Castillo, C., Donato, D., Leonardi, S., & Millozii, S. 2006. Temporal Analysis of the Wikigraph. In *Web Intelligence*, 45-51. Hong Kong, China: IEEE CS Press.

Cherones, T. 1992. *The Limo*. Seinfeld. J. and David, L. United States.

Denning, P., Horning, J., Parnas, D., & Weinstein, L. 2005. Inside Risks: Wikipedia Risks. *Communications of the ACM* 48(12): 152.

Gretzel, U. and Fesenmaier, D. R. 2002. Building Narrative Logic into Tourism Information Systems. *IEEE Intelligent Systems* November/December, 2002: 59-61.

Hart, P. E., Nilsson, N. J., & Raphael, B. 1968. A Formal Basis for the Heuristic Determination of Paths in Graphs. *IEEE Transactions on Systems Science and Cybernetics* 4(2): 100-107.

Isbister, K. and Doyle, P. 2003. Web Guide Agents: Narrative Context with Character. *Narrative Intelligence*, 229-243. M. Mateas and P. Sengers eds. Philadelphia, Penn.: John Benjamins Publishing Company.

Kim, J.-W., Kim, C.-S., Gautam, A., & Lee, Y. 2005. Location-based Tour Guide Systems Using Mobile GIS and Web Crawling. *Web and Wireless Geographical Information Systems*, 51-63. Berlin, Germany: Springer**.**

Lanegran, D. 2005. Discussion on question, "What makes a good field trip?" With B. Hecht. St. Paul, Minn.

Mateas, M. and Sengers, P. 1999. Introduction to the Narrative Intelligence Symposium. *Fall AAAI Symposium on Narrative Intelligence*. Cape Cod, Mass.

Mateas, M. and Sengers, P. 2003. Narrative Intelligence (Introduction). *Narrative Intelligence*, 1-25. M. Mateas and P. Sengers eds. Philadelphia, Penn.: John Benjamins Publishing Company.

Mott, B., Callaway, C., Zettlemoyer, L., Lee, S., & Lester, J. 1999. Towards Narrative-Centered Learning Environments. *Fall AAAI Symposium on Narrative Intelligence*. Cape Cod, Mass.

Persson, P., Höök, K., & Sjölinder, M. 2003. Agneta & Frida: Merging Web and Narrative. *Narrative Intelligence,* 245-258. M. Mateas and P. Sengers eds. Philadelphia, Penn.: John Benjamins Publishing Company.

Propp, V. J. 1928. *Morphology of the Folktale*. ed. 1968, Austin, Tex.: University of Texas Press.

Schöning, J., Hecht, B., Rohs, M., & Starosielski, N. 2007. WikEar – Automatically Generated Location-Based Audio Stories between Public City Maps. In *9th International Conference on Ubiquitous Computing Demo Proceedings*, Innsbruck, Austria.

Wells, C. G. 1986. *The Meaning Makers: Children Learning Language and Using Language to Learn*. Portsmouth, N.H.: Heinemann.

Zachte, E. 2006. *Wikipedia Statistics*. Retrieved May 1, 2007, from http://stats.wikimedia.org/EN/TablesWikipediaEN.htm.