Facilitate Knowledge Exploration with Storytelling

Mei Si

Rensselaer Polytechnic Institute, Troy, NY
sim@rpi.edu

Abstract

With the fast development of natural language processing and web technologies in recent years, the amount of knowledge/data available for everyone to explore has grown exponentially. This project is aimed at creating an interactive and proactive presentation agent. In particular, narrative and storytelling techniques are used for engaging the audience and helping the audience digest and remember the content. We present examples of using this system for generating presentations, and preliminary evaluation results.

Keywords: Storytelling, information exploration

1 Introduction and Motivation

Narrative has always been an important part of human communication. We tell stories not only for entertainment, but also for sharing information and influencing others. Many researchers have argued that narrative is essential for how people understand and organize their experience [9, 11, 4, 1]. For example, Bruner argued, “We organize our experience and our memory of human happenings mainly in the form of narrative” [4]. Similarly, Neumann and Nünning believed “Narratives are not only a literary art form but a fundamental way of organizing human experience and knowledge,” and that narrative can be regarded as a “fundamental way of world making” [10]. Hermann suggested that there are “five ways stories scaffold intelligent behaviors” including chunking experience, imputing causal relations, problem raising/solving, sequencing actions, and distributing intelligence. Abbott even suggested that “narrative is the principal way in which our species organizes its understanding of time.” [1] There is also evidence from empirical studies that when communicating or presenting information in a narrative form, the relationships among the events and objects become clearer and therefore more understandable and memorable to others [14, 13, 6].

In this work, we propose a narrative-based tool for helping people explore information. With the fast development of natural language processing and web technologies in recent years, the amount of knowledge/data available for everyone to explore has grown exponentially. For example, there are more than four million entries currently in the English Wikipedia. Existing means for consuming such large data are often either non-interactive, e.g. summary reports, or passive, e.g. data query or visualization tools where the user needs to specify the queries.
The goal of this project is to create an automated agent for helping people explore large networks of information. The agent acts as a narrator for the knowledge network. Just like a human presenter, the agent picks and presents topics from the knowledge base one by one. The user can interrupt and change the topic. The user can also influence how the agent selects future topics. While deciding what to say, the agent strives to provide the user a structured experience by proactively constructing narratives using information relevant to the user’s interests and balancing a number of objectives, such as topic consistency and novelty, tension management, and various local and global narrative structures. At a more local level, the agent also illustrates and emphasizes the relationships among adjacent topics by constructing transition sentences using analogies and contrasts.

Figure 1 shows the overall pipeline of how the presentation agent works. It takes information as a network of topics which may come from a variety of sources, including social media, crowdsourcing, and querying existing web knowledge base such as DBpedia. The presentation agent performs narrative planning as it is interacting with the user. We anticipate different interactive visualization tools will be needed for different application domains and have taken steps to build two of them.

Figure 1: System Pipeline

Many attempts have been made to design virtual characters that can make presentations, such as [5, 7, 12, 2, 3]. Most existing works have strict requirements for how the domain knowledge should be encoded. As a result, the domain knowledge needs to be designed by hand which is severely time-consuming. In this work, we experiment with loosening up such constraints and making minimal assumptions for the type of information the agent gets for enabling the agent to generate presentations for a wider range of domains. The presentation agent also has an objective of encouraging the user to explore the data rather than just making a presentation.

2 Example Domain and Knowledge Representation

Throughout this paper, we will illustrate the working of our systems by introducing the 2008 Summer Olympic Games in Beijing, China.

For encoding the agent’s knowledge about a domain, we hope to use a knowledge representation that is both compatible with structured data, such as results from querying DBpedia, and is intuitive enough for non-technical authors to design and edit the knowledge base manually. As an initial step, we created a XML format that encodes knowledge as a directed graph. Each topic the agent can talk about is represented as a node with a unique ID, and nodes are linked to each other by their relationships. For example, the Beijing Olympic Village is located in Olympic Green, which is one of the Olympic parks. In the XML, both “Beijing Olympic Village” and “Olympic parks” are nodes, and “is in” is the relationship between them.

In addition to the relationship between them, both nodes have multiple links to other nodes in the knowledge base. Each node may also contain multiple tags for additional information.
Each tag has a tag type and tag value. Currently, we are only using the “FunFact” tag for storing additional funny facts about the node.

Overall, the graph of knowledge looks very similar to a concept net [8]. Instead of each piece of knowledge being a concept, here we allow the nodes to represent any information. Currently, the XML file for this domain includes 300 nodes and more than 1000 relationships.

3 Narrative Agent

We hope to design an automated agent that can act as a personal assist/guide for introducing the information in a knowledge base. When deciding what to talk about next, the agent strives to form a piece of narrative rather than simply listing the facts. It does so in two steps. First, it plans the sequence of topics to be presented. In this process, it considers and balances multiple narrative and user experience related objectives, such as suggesting novel content from time to time, following an organizational or geographical order in describing the topics, and leaning towards topics that the user may be interested in. Secondly, when entering a new topic, it addresses the topic’s relationships with previous topics. The new topic may be a continuation of the previous topic, be analogous to a topic discussed before, or bring the conversation to a new direction. Though explicitly pointing out these relationships, the agent helps the user to establish connections among the content, and to anticipate what will be presented next.

3.1 Topic Planning by Multi-Objective Optimization

When deciding what to talk about next, the agent takes a number of objectives into consideration, including forming appropriate local and global structures in the presentation, addressing the user’s interests, and maintaining the user’s engagement level. Each objective is weighted with a relative importance. The agent picks its next topic node by maximizing its satisfaction of all the objectives with their relative importance factored in. Currently, we only consider hierarchical ordering consistency and spatial ordering consistency for forming the local structure in the description. By default, we give high importance to the objectives of maintaining these ordering consistencies. The user can directly change the style of how the agent picks the next topic by modifying the relative weights of the agent’s objectives. Thus, the agent both has its own agenda, and is reactive to the user’s interests and preferences.

Algorithms 1-4 provide the pseudo codes for deciding how much picking a particular node will contribute to each of these objectives. At each turn of the interaction, the narrative agent
computes a weighted sum of each node’s satisfaction levels to these objectives, and the node with the highest score will be picked. Next, we will briefly describe how these objectives are evaluated.

Algorithm 1 Hierarchical_Ordering_Consistency (node, current)

#node: The potential next node  
#current: The current node  

keywords ← [“contain”, “include”]  #Keywords for identifying hierarchical relationships  
neighbor ← Get_Common_Neighbor(node, current)  

if neighbor then  
    relationship1 ← Get_Relationship(node, neighbor)  
    relationship2 ← Get_Relationship(current, neighbor)  
    if relationship1 == relationship2 && relationship1 in keywords then  
        return 1  
    return 0

For knowledge that is structured hierarchically, the agent will talk about items at the same hierarchical level together. Algorithm 1 allows the agent to check whether the potential next topic is at the same hierarchical level as the current one. Similarly, when describing a physical environment, the agent will describe the items in it in order, e.g. clockwise or counter-clockwise. Algorithm 2 allows the agent to check whether that order is preserved.

Our system supports the user to interact using speech, typing, and mouse clicking. To take the user’s interest into consideration, the agent remembers the last concept mentioned by the user, and favors all the related concepts in subsequent dialogue (Algorithm 3). Currently, a simple breadth-first search (BFS) is used for computing the distance between any pair of nodes in the network. The shorter the distance, the more coherent the node is to the user’s interest.

Algorithm 2 Spatial_Ordering_Consistency (node, current, previous)

#node: The potential next node  
#current: The current node  
#previous: The node narrated before the current node  

keywords ← [“totheright”, “totheleft”]  #Keywords for identifying spatial relationships  
relationship1 ← Get_Relationship(node, current)  
relationship2 ← Get_Relationship(current, previous)  

if relationship1 == relationship2 && relationship1 in keywords then  
    return 1  
return 0

Finally, the agent wants to keep the user engaged through talking about new and interesting information (Algorithm 4). Since we cannot automatically evaluate how interesting a piece of node description is, the designers of the XML files need to use the ”FunFact” tag to signal them. To estimate how novel a piece of information is to the user, the agent considers three factors: whether the node has been talked about before (Addressed(node)), what percentage of its immediate neighbors have been talked about (Neighbor_Addressed(node)), and how far the node is from the current topic. These three factors can be weighted. Typically, if a node has already been presented, it is considered to have very low novelty. Similarly, the relative importance of the node being novel and containing an interesting description can be weighted by the author.
### Algorithm 3 Adhere To User Interest (node, interest, maxDis)

#node: The potential next node
#interest: The last node queried by the user in the past 5 steps
#maxDis: The longest distance between two nodes in the knowledge base

if interest then
    distance ← BFS(node, interest)
return 1 - distance/maxDis
return 0

### Algorithm 4 Engagement (node, current)

#node: The potential next node
#current: The current node
#funFact: Whether this node contains a FunFact
#maxDis: The longest distance between two nodes in the knowledge base

distance ← BFS(node, current)/maxDis
novelty ← distance*\(w_0\) - Addressed(node)*\(w_1\) - Neighbor Addressed(node)*\(w_2\)
return novelty*\(w_3\) + funFact*\(w_4\)

### 3.2 Connecting Topics by Making Analogies and Contrasts

The descriptions about individual topics do not include its relationship to other topics. To reveal and emphasize such relational information, the narrative agent applies three techniques: making an analogy between the next topic and something that has been mentioned before; explicitly mentioning how the next topic is related to the current one; and signaling a topic transition. If none of these apply, the agent will use a simple connection phrase, such as “Next” or “There is also” to lead to the next topic. Algorithm 5 describes this process.

### Algorithm 5 Link Topics (node, current)

#node: The next node
#current: The current node
#history: All the nodes that have been presented before
#keywords: Relationships for making analogies

relationship1 ← Get Relationship(node, current)
if relationship1 then
    for n in history do
        #Get the relationship between two consequent topics
        relationship2 ← Get Relationship(n, n.next())
        if relationship1 == relationship2 && relationship1 in keywords then
            return Make Analogy(node, current, n.next(), n)
    return Mention Relationship(node, current, relationship1)

distance ← BFS(node, current)
if distance > threshold then
    return Topic Transition
else
    return Non Topic Transition
Analogies are preferred over other lead-in sentences because it can not only provide ground for the next topic but also refresh the user’s memory about the previous topics. To draw an analogy, the agent compares the most recent node transition to each previous transition. If the same relationship appears in both transitions, an analogy is added to the current node’s information when it is presented. To decrease the appeared repetitiveness, the structure of the analogy sentence is chosen at random from a set of templates. For example, the agent may say “Just as Fencing includes (Fencing) Women’s Sabre, so too Tennis includes (Tennis) Women’s Doubles.”

Lead-in sentences are based on the most recent transition. If there is an explicitly labeled relationship between the two most recent nodes, then it is stated. Otherwise, depending on how far away the next topic is from the current one, the agent will either signal a topic transition or use a simple connection phrase. Similar to making analogies, the structure of the lead-in sentence is chosen at random from a set of templates. A special case is when the agent may start with “first off” for its very first sentence of the presentation. If the agent knows the length of the presentation, it may add “finally” to its last sentence. The following paragraph demonstrates some examples of using analogies and lead-in sentences. The same paragraph was used in our evaluation without the annotations.

First off, let’s talk about Angela Hucles. (Add lead-in sentence for opening a new topic)

Angela Hucles is a retired American professional soccer midfielder and was a member of the United States Women’s National Soccer Team. (State node description)

Angela Hucles won a gold medal in Women’s Football Tournament. (Point out the relationship between the current and next topics)

The women’s association football tournament at the 2008 Summer Olympics took place in Beijing and four other cities in China on 6-21 August. (State node description)

Now then, about Football. (Signal topic switch)

LaShawn Merritt also won a gold medal in Men’s Track and Field 4x400 m Relay, similar to how Angela Hucles won a gold medal in Women’s Football Tournament. (Make analogy)

4 Evaluation

We believe this system can aid people in information exploration in three ways: providing an interactive visualization shell; presenting topics in a meaningful and interesting sequence; and illustrating the relationships among the topics introduced. A large volume of research already exists on the human factor issues related to data visualization and human-computer interaction. In this work, we want to focus on our narrative agent. We performed an empirical study to evaluate the effect of the last item – illustrating the relationships among the topics – on people’s experience of the presentation. We hypothesize that people will like to see presentations with these sentences added by the agent.

The evaluation was conducted by recruiting subjects from Amazon mTurk. Instead of using the visualization frontend, we presented the agent’s constructed narrative in plain text to the subjects and let the subjects rate how clear, coherent, and comprehensible the presentation is. These metrics are adapted from [13]. A 7-point Likert scale was used for the ratings, where 1 stands for not clear/coherent/comprehensible at all, and 7 stands for completely clear/coherent/comprehensible. The experiment utilized two presentations. Each contains a sequence of eleven topics. Each sequence was picked by the narrative agent’s topic selection algorithm without using look ahead. One presentation is about Angela Hucles as shown above; the other presentation is about the sport of fencing.
Based on the topic sequences and topic descriptions, we created three versions of the presentations. The first version includes only the topic descriptions. The second version includes the topic descriptions and the relationship descriptions when two consecutive topics are directly related. Finally, the third version was generated using Algorithm 5, and the agent prefers to make analogies rather than describing the relationships. Other lead-in sentences were also used for either signaling a topic switch or continuing the topic. We expected the subjects to like version 2 and version 3 better than version 1.

We recruited 35 subjects for rating the presentations about Angela Hucles, and 45 subjects for rating the presentations about fencing. A within-group design is adopted where each subject read and rated all three presentations. The mean ratings are shown in Figure 2.

We found a similar pattern in the subjects’ ratings of the presentations. For both presentations, the ratings of version 1 and version 3 are very similar, and the ratings of versions 2 are much lower. 2-tail paired sample t-tests were performed for comparing the ratings of version 1 and version 3 to those of version 2. For the presentations about Angela Hucles, all the t-tests are significant at .05 level. For the presentations about fencing, the t-tests on the differences of ratings for clarity and coherence are all significant at .05 level. Both t-tests for comparing the ratings for comprehensibility are not significant. We also compared the ratings between version 1 and version 3. None of the comparisons is significant.

These results suggest that either not adding relationship descriptions, or adding analogies and lead-in sentences between each pair of topics is more welcomed than inserting the relationship descriptions alone. This is different from our hypothesis. We expected the subjects to like version 2 better than version 1 because version 2 provides more context – the relationships – for understanding and remembering the information. The subjects were encouraged to leave explanations for their ratings. There is no direct feedback from the subjects on why version 2 is preferred least. Some subjects mentioned that version 3 is less dry and flows better. For both presentations, about 1/3 of the successive topic pairs has a direct relationship. We suspect that adding the relationship descriptions makes the text longer to read and seems less structured, which leads the subjects to like them less than version 1. The presentations are even longer in version 3. However, because the subjects can see the transitions of topics more clearly, understanding the structural of text may feel much easier.

![Figure 2: Mean Ratings for a) Angela Hucles, and b) Fencing](image)

5 Discussion and Future Work

Narrative and dialogue are good ways for engaging people and helping them organize and memorize information. This project is aimed at helping people consume information in a narrative form. The system has been fully implemented. We plan to conduct more formal evaluations on how the narrative agent or the visualization interface can help improve user’s experience, engage the user in exploring the data, and increase the user’s recall of the information being presented.
Aside from evaluation, we want to improve the algorithms for evaluating the objectives. Firstly, we calculate the distance between two nodes simply using BFS over the knowledge graph. This distance may not correspond to the distance between the two concepts in the user’s mind. In the future, we are considering using a semantic tool such as concept net [8] to help us get a better estimation. Secondly, both the evaluations of hierarchical ordering constraints and spatial ordering constraints require there be no gap in following the orders. The agent thus cannot talk about something that is irrelevant to the hierarchical or spatial order it is trying to follow, and then come back. Finally, we want to support a richer set of narrative techniques for making the presentation. Currently, the agent plans its narrative as a single thread through the knowledgebase. In movies and novels, often the time multiple interweaving storylines are presented. We are going to explore this technique for creating a more interesting and engaging experience for the user. We are also interested in using foreshowing and flashback as ways for scaffolding the user as well as modulating the user’s excitement and emotional experiences in the data exploration process.

References