ABSTRACT
This paper describes a digital theater installation call Buzz. Buzz consists of virtual actors who express the collective voice generated by blogs. These actors find compelling stories from blogs on the popular and contentious topics of the day and perform them. In this paper, we explore what it means for a story to be compelling and describe a set of techniques for retrieving compelling stories on a topic. We also outline an architecture for high level direction of a performance using Adaptive Retrieval Charts (ARCs), allowing a director level of interaction with the performance system. Our overall goal in this work is to build a model of human behavior on a new foundation of query formation, information retrieval and filtering.

Keywords
Network Arts, Emotion, Blogs, Media Arts, Culture, World Wide Web, Software Agents, Story Generation

1. INTRODUCTION
The Internet is a living, breathing reflection of who we are, what we think, and how we feel. The pages that make up the Web form the book of our contemporary life and culture. They are the ongoing and changing buzz of our world.

The latest embodiment of this cultural reflection is found in blogs. Vast numbers of people look to blogs as a means of hearing about current events, communicating with friends, meeting people with similar interests, forming communities, sharing media, and expressing their ideas on their personal soap box. Not only are blogs widespread, but they are incredibly dynamic, with hundreds updated per minute. The existence of millions of blogs on the web has resulted in more than the mere presence of millions of online journals, but rather, they generate a collective buzz around the events in the world.

Buzz (shown in Figure 1) is a multimedia installation that exposes the collective voice generated by blogs. Using the most popular searches of the moment, it finds the blogs that are reflections of the questions of our time. The most popular searches are displayed on a central screen. The blogger is embodied by virtual actors, displayed on surrounding displays, who externalize these monologues by reading them aloud. It exposes the opinions and experiences of people in the world relating to what we are searching for.

As an example of a Buzz performance, Table 1 shows the buzz on the topic of the popular reality television show “Project Runway.” In the performance, the words “Project Runway” appear on the central screen, while the four actors contribute to the performance by reading these three discovered stories (found in blogs) aloud, in turn. The actors are attentive to each other by turning to face the actor currently speaking.

I never watched Project Runway in its first season, and only after the goading of a few people did I consent to test out the new season. It’s surprisingly good, and by ”good” I mean ”honestly quality television,” not “guilty-pleasure-o-rama.” What distinguishes it from other reality shows, to me, is that the contestants are professionals doing what they do best.

Bravo, why do you show project runway reruns every week before the real project runway is on? I hate you. I hate you. I don’t want to look at Santino’s bad underwear again, I want to see Pretties. UGLY. UGLY. And Daniel 2 had such a great design concept, but the execution was BLAH.

However, as a longtime television viewer, I think that I’m fairly qualified to know what is “good” and “bad” tv. Project Runway is more than just a guilty pleasure. I agree with Victor Balta’s assessment that there’s a greater focus on talent in this show.

Table 1: Three stories discovered by Buzz, on the topic “Project Runway”, a popular reality television show.
2. THE ASSOCIATION ENGINE

This work began as a project called the Association Engine [13], composed of a troupe of virtual improvisational actors. A troupe of five actors, with animated faces [15] and voice generation [11], began a performance by taking a single word or phrase suggestion from the audience, through keyboard input. They used this word as a seed to an improvisational warm-up game called the Pattern Game, where the actors free associate to create a collective context, getting themselves on the same contextual page.

Following this warm-up game, the actors would generate a One Word Story, from the context of the warm-up. A One Word Story is a common game in improvisational theater where actors each contribute one word at a time to create a collective story. See Table 2 for a sample pattern game and generated One Word Story from the Association Engine.

3. RELATED WORK

In building a story generator for the Association Engine, we faced problems that prevailed from previous years of Artificial Intelligence research in story generation. TaleSpin [10] used a world simulation model and planning approach for story generation. To generate stories, TaleSpin triggered one of the characters with a goal and used natural language generation to narrate the plan for reaching that goal. The stories were simplistic in their content (using a limited amount of encoded knowledge) as well as their natural language generation.
Klein’s Automatic Novel Writer [6] uses a similar approach in order to produce murder novels from a set of locations, characters, and motivating personality qualities. The stories follow a planning system’s output as the characters searched for clues. The system does not seem to capture the qualities of a good murder story, including plot twists, foreshadowing, and erroneous clues.

Dehn’s Author system was driven by the need for an “author view” to tell stories, as opposed to the “character view” found in world simulation models. Dehn’s explanation was that “the author’s purpose is to make up a good story” whereas the “character’s purpose is to have things go well for himself and for those he cares about” [2].

Using a template-based approach, the Association Engine was able to generate stories that were coherent, but did not engage the audience, as seen from the sample One Word Story in Table 2. They lacked in character development and a general purpose.

In general, previous story generation systems faced a trade-off between a scalable system, and one that can generate coherent stories. Besides Dehn’s Author, previous research in this area has employed a weak model of the aesthetic constraints of story telling.

In response to the shortcomings of story generation, we explored story discovery. We found an incredible corpus of existing stories of people’s life experiences. These stories exist within a subset of blogs [8, 4] found on the Internet. We then changed our focus from story generation to story discovery, what people are currently thinking and saying online. We define a stronger model for the aesthetic elements of story telling and use that to drive retrieval of stories, and to filter and evaluate results.

Artistically, story telling and online communication have been externalized within several installations. Of the more well-known, Listening Post [5] exposes content from thousands of chat rooms through an audio and visual display. Similar to Listening Post, Buzz externalizes online communication, though in a more contextually grounded manner. Like Buzz, Mateas’s Terminal Time [9] sought to tell stories, though its stories were grounded in a common sense knowledge base as opposed to live real world information.

4. TELLING COMPELLING STORIES
Starting with the “Pattern Game” as a context, we enabled Buzz to discover stories by searching for blogs that related to the terms resulting from “Pattern Game.” A first pass at building Buzz revealed that the content of blogs is incredibly wide ranging, but unfortunately often very dull. Buzz succeeded in finding stories that were on point to the current context, but the results were not compelling.

We found that people blogged about topics including their class schedule, what they are eating for lunch, how to install a wireless router, what they wore today, a list of their 45 favorite ice cream flavors. While this was interesting to observe from a sociological point of view, it did not make for a compelling performance. Not only were the blogs on these topics boring, but the lengths of the stories varied widely from one sentence to pages upon pages.

We needed to give the system strategies for finding stories that were compelling and engaging to an audience. To do so, we define a simple model for the aesthetic qualities of a compelling story. These qualities include but are not limited to:

- an interesting topic
- emotionally charged
- complete and of a length that holds the audience’s attention
- at the right level of familiarity (not overly common and not too obscure)
- involving dramatic situations

We designed Buzz to find stories with all of these qualities.

4.1 Topics Of Interest
A compelling story is generally about a compelling topic, one that interests the audience. For this reason, we chose the day’s most popular searches from Yahoo (provided by Yahoo buzz [20]) as topics. Search engines recently started providing a log of their most frequently used query topics. This feed worked well as a seed to story discovery, as we are using the topics that people are searching for most, and discovering people’s thoughts and opinions on these topics.

We found Wikipedia [19] to be another source for topics of interest as the site maintains a list of “controversial topics”. The list shows topics that are in “edit wars” on Wikipedia as Wikipedia contributors are unable to agree on the subject matter. This list includes topics such as apartheid, overpopulation, ozone depletion, and censorship. These topics, by their nature, are topics that people are passionate about.

Using these two sites as sources for topics, finding compelling stories began with a simple web search restricted to the domain of www.livejournal.com [8], with each focal topic as a search query. Out of the first 100 results for each topic, about 60 tend to be actual blog entries and not blogger’s profile pages (this differs greatly per topic).

After discarding profile pages, the remaining blog entries are analyzed phrasally, eliminating posts that do not contain at least one of the two word phrases (non-stopwords) from the topic. For example, given a topic of ‘Star Wars: Revenge of the Sith,’ entries that contained the phrase ‘star wars’ were acceptable, but not entries that merely had the word ‘star’ or ‘wars.’ The remaining blog entries were known to be relevant to the current popular topic.

Using topics of interest as the source of topic keywords and blogs as the target, we were able to generate what was being said about what people were most interested in.

4.2 Filtering Retrieval by Affect
Given that our initial version of Buzz was reading blogs that were not compelling, and since such a large volume of blogs exist on the web, we strove to filter the retrieved blog entries by affect, giving us the ability to portray the strongest affective stories. Beyond purely showing the strongest affective stories, we also wanted to be able to juxtapose happy stories on a topic with angry or fearful stories on a topic.

To build such a tool, we sought a large corpus of affectively scored terms, and we found one. The ANEW [1] corpus contains 1,034 unique words with affective valence (a scale from unpleasant to
pleasant, and dominance (a scale from submissive to dominated) with scores on a scale of 1-9. Using the ANEW as a base, we expanded the corpus by synonyms from Thesaurus.com [18]. As some synonyms are more closely related to a word than others, we used web co-occurrence to determine how closely related terms are.

Let $W(s)$ represent a function that returns the web document frequency for a given query string $s$:

$$W(\text{Synonym} + \text{Word}) \quad W(\text{Word}) + W(\text{Synonym}) \quad (1)$$

yields a measure of the web co-occurrence of the two terms. Higher co-occurrence indicated that words are more closely related, and the synonym expansion reflected this relationship.

We created a tool to compile a collective score for a group of words based on the presence of keywords from our expanded corpus [16]. In addition to simple keyword detection, the tool also used negation detection to derive emotional content. For example, if someone says, “I am not happy,” straight keyword detection would fail to reveal the true emotional state of the writer.

The valence, arousal and dominance model of emotion did not give us the affective precision that we desired. To solve this, we used a simple mapping of the scores from the VAD (Valence Arousal Dominance) model to Ekman’s six emotion model (happiness, sadness, anger, disgust, fear, surprise) [7, 3]. We were then able to classify each blog entry as happy, sad, fearful, disgusting, angry, surprising, or neutral.

When using the emotional filtering tool, Buzz was considerably more compelling. The actors were then able to retrieve stories from the Web based on emotional stance, enabling the theatrical agents to juxtapose happy and angry stories on the same topic.

Further research in emotional classification has allowed us to build a more robust tool for general purpose affective classification [14, 7, 12]. We are currently developing new techniques for general purpose emotional classification which will increase our accuracy in filtering retrieval by emotion.

### 4.3 Filtering Retrieval by Syntax

In our first pass at retrieving stories from blogs, we noticed that we often found lists or surveys instead of text in paragraph form. For example, one blogger posted an exhaustive list of lip balm flavors. Others posted answers to a survey about themselves (their favorite vacation spot, favorite color, favorite band and actor, etc.). These are clearly not good candidates for compelling stories.

To solve this problem, we chose to filter the retrieved blog entries by syntax. Blog entries that met any of the following criteria were filtered:

1. too many newline characters (more than six in a entry of four hundred characters)
2. too many commas (more than three in a sentence)
3. too many numbers (more than one number in a sentence)

This method successfully filtered blog entries that contained a list or survey of some sort. While the precision of such removal of blogs based on syntax was lower, we optimized for recall so that all potential lists and surveys were removed for the corpus. Given the large volume of blogs on the web updated every minute, letting some potentially good blogs fall through the cracks sufficed for our purposes.

### 4.4 Colloquial Filtering

In order for an audience to find a story engaging, they must understand the story. That is, the story can’t involve topics that the audience is unfamiliar with or contain jargon particular to some field. The story must be colloquial.

To determine how colloquial a story is, we built a classifier that makes use of page frequencies on the web. For each word in the story, we look at the number of pages in which this word appears on the web, a frequency that is obtained through a simple web search. Applying Zipf’s Law [21], we can determine how colloquial each word is [17]. A story is then classified to be as colloquial as the language used in it.

### 4.5 Complete Passages

Given the blog entries that remained after passing through the four above mentioned filters (relevance, affect, syntax and colloquial), the agency must choose which pieces of blog entries to present to the audience. This involves finding complete thoughts or stories of a length that can keep the audience engaged.

For the most part, we found that blog authors format their entries in a way such that each paragraph contains one distinct thought. Given this trend, the paragraph where the topic of interest is mentioned with the greatest frequency will suffice as a complete story for our system. If this paragraph is of an ideal length (between a minimum and maximum threshold), which we determined by viewing Buzz with stories at many different lengths, then it is posted as a candidate story. For our system, we found that stories between 150 and 400 characters long were ideal. Again, given the large volume of blogs on the web, letting many blogs fall through the cracks because they are too long or too short, is fine for our purposes.

An example of three stories discovered by Buzz on the topic of “Project Runway,” a popular reality television show, can be seen in Table 1. The stories shown were retrieved and passed through all above mentioned filters. Notice the differing emotional stances of the first and second stories. This was a deliberate juxtaposition of positive and negative passages on this topic.

### 5. DIRECTOR LEVEL CONTROL

Given the above classifiers and filters, we are able to retrieve a set of compelling candidate stories on a particular topic. These filters and classifiers also give us a level of control of the performance similar to that of a director. Having information about each story such as its “emotional point of view”, and its “obscenity”, we can plan out the structure of the performance from a high level view, giving the performance a flow, based not only on content, but on emotion, obscurity, on-point vs. tangential, etc. Given a topic, we can juxtapose stories with different emotional stances, different levels of obscurity, and on-point vs. off-point. These affordances give a meaningful structure to the performance.

To provide a high level control of the performance, we created an architecture for driving the retrieval of performance content.
Figure 2: A sample ARC from the Buzz system, defining a point/counter point interaction between agents.

The structures, called Adaptive Retrieval Charts (or ARCs), provide high level instructions to the buzz engine as to what is needed, where to find it, how to find it, how to evaluate it, how to modify queries if needed and how to adapt the results to fit the current goal set.

An example of an ARC used in Buzz is shown in Figure 2. The pictured ARC defines a point/counterpoint/dream interaction between agents. The three modules define three different information needs, as well as the sources for retrieval to fulfill these needs. The first module specifies that we want a blog entry that is on point to a specified topic, has passed through the syntax and colloquial filters, and is generally happy on the topic. The module specifies using Google Blog Search [4] as a source. The source node specifies to form queries by single words as well as phrases related to the topic. If too few results are returned from this source, we have specified that queries are to be continually modified by lexical expansion and stemming.

The ARC extensible framework allows for interactions from directors with no knowledge of the underlying system. In a future system, we will accomplish this via a range of possible interfaces from storyboarding and affect manipulation to a natural language interface.

6. BUZZ IN THE WORLD

Enabling Buzz with the ability to discover compelling stories on a popular topic has produced great results. Buzz has changed from an installation that was unbearably dull, exposing the boring nature of many blogs, to a system that engages its viewers. The performance is now not driven simply by the relevance of on-line content, but by the blogger’s emotional state. The highly emotional content engages the audience and affords us a high visibility installation.

Buzz was exhibited last year at the Athenaeum Theater as a part of the 8th Annual Chicago Improv Festival. It was received with great excitement. Actors, writers, producers and theater goers alike viewed Buzz in this 10 day installation.

“We are excited about having Buzz at the Chicago Improv Festival because, like improvisers, it goes beyond the events of the day, to a point of view,” commented Jonathan Pitts, Co-Founder and Executive Producer of the Chicago Improv Festival.

Buzz was displayed in the lobby of Chicago’s Second City theater at 1616 N. Wells St. in Chicago on August 24th, 2005 for a long term installation, currently still running.

7. DRAMATIC SITUATIONS

Through experiencing Buzz in the world and watching audiences reactions and responses to stories, we discovered more generalized traits of compelling stories. The most compelling stories to watch were those where someone is laying their feelings on the table, exposing a dream or a nightmare that they had, making a confession or apology to a close friend, or regretting an argument that they had with their mother or spouse.

Codifying these qualities, we built a story discovery engine to seek out these types of stories. While still making use of multiple retrieval filters described in the previous section, we added a component to the retrieval that found stories that began with a cue that the writer was about to describe a dream, nightmare, fight, apology, confession, or any other emotionally fraught situation. Such cues include phrases such as “I had a dream last night,” “I must confess,” “I had a terrible fight,” “I feel awful,” “I’m so happy that,” and “I’m so sorry.”

This realization was an important turning point in our system’s capabilities with regard to retrieving compelling stories. The newest instance of Buzz no longer focuses on the popular or contentious topics, but instead focuses on stories in different types of emotion-laden situations (dreams, fights, confessions, etc.). See Table 3 for examples of stories found by this system.

These stories are more interesting as the blogger isn’t talking about a popular product on the market, or ranting about a movie, they are relating a personal experience from their life, which typically makes them emotionally charged. The experiences they describe are often frightening, funny, touching, or surprising. They describe situations which have a common element in all of our lives, giv-
I have a confession – beneath my cynical, sarcastic facade beats a heart of pure mush. Before you snort milk through your nose, think about it – despite Connie’s best efforts, my favorite movie in the world is THE SOUND OF MUSIC. What’s not to like? Great songs, fabulous scenery, incorrigible children, a charming nun/governess and a stern, handsome frozen-hearted captain who slowly melts under the spell of the songs, the scenery, his kids and Julie Andrews. When I start that movie and the mountain scenery comes on the scene with the birds twittering and the first chords of music play … I’m in heaven.

My husband and I got into a fight on saturday night, he was drinking and neglectful, and I was feeling tired and pregnant and needy. It’s easy to understand how that combination could escalate, and it ended with hugs and sorries, but now I’m feeling fragile. Like I need more love than I’m getting, like I want to be hugged tight for a few hours straight and right now, like I want a dozen roses for no reason, like a vulnerable little kid without a safety blankie. Fragile and little and I’m not eating the crusts on my sandwich because they’re yucky. I want to pout and stomp until I get attention and somebody buys me a toy to make it all better.

Last night for instance, I dreamed that we were having the rehearsal dinner at an aquarium for some reason this aquarium had a killer whale and I was dumb enough to dip my feet in the tank. Well, it attacked, and in the dream I was clearly bummed out due to having a major foot surgery instead of a wedding. There was also a debacle with a scorpion that I won’t go into. And also the cake melted.

| Table 3: Three stories discovered by Buzz. |

8. FUTURE WORK
Initially, story discovery within Buzz was based on popular topics. As we approached the task of engaging the user, it became more important that the stories themselves were compelling, as opposed to topical. Using filters and information retrieval strategies that focussed on finding the interesting and not the topical has resulted in an engaging theatrical installation. In the future, we will turn our focus back to topics, discovered within the scope of interesting stories.

Our current and future work in this area involves expanding Buzz into a full length improvisational performance on stage, interacting with human actors. We are building a full body projected avatar host with voice generation, and voice recognition to take audience suggestions and interact with human actors. Understanding the state of current technology in voice recognition, we are enabling the host to drive her conversations with actors and the audience, to recover from mistakes, and express and expose her shortcomings.

This production will make use of the ARC architecture to allow a high level control of the flow of the performance. Our research in story discovery will serve as a platform for character development for the host, as she can relate to and participate in discussions by telling stories discovered from blogs related to the current conversation topic or audience suggestion. This will also involve expanding our set of available filters to include a text based gender (of the author) classification tool.

9. REFERENCES


