AAAI Technical Report WS-11-18

# Towards a Computational Model of Narrative Visualization

**Alok Baikadi  Julius Goth  Christopher M. Mitchell  Eun Y. Ha  Bradford W. Mott  James C. Lester**
Department of Computer Science, North Carolina State University, Raleigh, NC, USA
{abaikad, jgoth, cmmitch2, eha, bwmott, lester}@ncsu.edu

**Abstract**

The task of narrative visualization has been the subject of increasing interest in recent years. Much like data visualization, narrative visualization offers users an informative and aesthetically pleasing perspective on "story data." Automatically creating visual representations of narratives poses significant computational challenges due to the complex affective and causal elements, among other things, that must be realized in visualizations. In addition, narratives that are composed by novice writers pose additional challenges due to the disfluencies stemming from ungrammatical text. In this paper, we introduce the NARRATIVE THEATRE, a narrative visualization system under development in our laboratory that generates narrative visualizations from middle school writers' text. The NARRATIVE THEATRE consists of a rich writing interface, a robust natural language processor, a narrative reasoner, and a storyboard generator. We discuss design issues bearing on narrative visualization, introduce the NARRATIVE THEATRE, and describe narrative corpora that have been collected to study narrative visualization. We conclude with a discussion of a narrative visualization research agenda.

## Introduction

The study of narrative has had a long history in artificial intelligence. Narrative plays a prominent role in cognition, and narrative understanding plays a foundational role in both understanding fictional worlds as well as the world around us. By supporting this role, narrative can serve as a focus for studying creativity, causal and temporal reasoning, and communication. This has led to systems that use narrative for learning environments in educational and training contexts.

While we often think of narratives as linguistic entities, there are many communicative modes in which a narrative can be delivered. Many forms of narrative have strong visual components, such as film, painting and photography. Others have more auditory components, such as music and the spoken word. These modes of storytelling, the art of presenting narratives, are each able to convey different types of information. In the past decade, there has been increasing interest in computational models of narrative that automatically translate between such modes.

*Narrative visualization* is the process of transforming narrative text into a visual space consisting of an environment populated by characters and objects. One of the major branches of research in natural language visualization (NL visualization) is the task of inferring the environment from a given text (Coyne et al. 2001; Glass and Bangay 2008; Zeng et al. 2005). Others (Joshi et al. 2006; Schwarz et al. 2010) have focused on using existing photo sources to illustrate a given text. There are a few systems (Johansson and Berglund 2005; Ye and Baldwin 2008) that impose some limitations on the input language in order to visualize events in a more narrative sense.

Narrative visualization poses several key challenges to a visualization system. In addition to the factual content, narratives also contain affective, causal, and temporal information that must be maintained. The fundamental unit of a narrative is an event, which describes how characters interact with the story world. Therefore, visualizing these events is a central problem of narrative visualization.

In this paper, we describe a computational framework that we have been exploring for narrative visualization. The framework leverages knowledge about the writing domain, along with semantic, temporal, and structural information in order to reason about the events in a narrative. It can then utilize this information to create a visual representation of each event.

Many approaches to narrative visualization require substantial knowledge engineering to define a visual semantics for the natural language text. Our approach reduces the amount of knowledge engineering required by targeting a specific domain of writing, and using machine learning techniques to translate between the written semantics and the visual semantics we have established for that domain.

Our work centers on a creativity support tool, the NARRATIVE THEATRE, which is being designed to aid developing writers in creating fables. There is near-consensus in the community that creativity requires the following distinct mental resources: intellectual abilities, knowledge, personality, motivation, and environment. In order to support knowledge and motivation, the NARRATIVE THEATRE presents the writer with a visual representation of her work. This visual representation aims to be similar enough to the written word for the feedback to support creativity in the revision process.

This paper is organized as follows: In the next section, we discuss previous work that has influenced our design decisions. Following this is a section presenting the key challenges and desiderata for a narrative visualization system for novice writers. We next present the system architecture for the NARRATIVE THEATRE, and the interactions with students that have been observed. We then turn to our design decisions with a summary of work to date, and conclude with a discussion of future work.

## Background

A central challenge of narrative understanding is reasoning over time. After discussing related work on temporal inference in natural language processing and considering how it applies to fable visualization, we briefly survey other issues in narrative visualization. Much of the work on temporal reasoning is based on the interval logic (Allen 1981), which is a general logic for reasoning about time in any context. The TimeML markup language (Pustejovsky et al. 2003) extends interval logic for natural language reasoning.

Jurafsky and Chambers (Chambers and Jurafsky 2008a; 2009) have utilized the TimeML system for temporal reasoning and have extended the representation for narrative event reasoning. The Narrative Event Chains model (Chambers and Jurafsky 2008b) extracts event information from the text and analyzes co-occurrence information, along with temporal reasoning results supported by TimeML, to discover likely sequences of narrative events. The Narrative Event Schemas model (Chambers and Jurafsky 2009) extends the Narrative Event Chains model by allowing for multiple protagonists and for general protagonist roles.

The SCHEHERAZADE system (D. Elson and K. McKeown 2009; D. K. Elson and K. R. McKeown 2007) presents a markup language to be used for learning. It distinguishes between narrative semantics, world knowledge, and story content. Their evaluation of SCHEHERAZADE revealed three major shortcomings. It was found that it was difficult to model beliefs involving past states, that in many cases details of the story were missing, and that users often wished for an "undo" button. In addition, the more expressive version of the model had much lower consistency between users.

With the recent appearance of the TimeBank corpus (Pustejovsky et al. 2003), data-driven approaches to modeling temporal relations have been gaining momentum. (Boguraev and Ando 2005) apply semi-supervised learning using a word profiling technique to recognize events and to infer temporal relations between time expressions and their anchored events in TimeBank. Mani et al. (2006) expand training data by computing the transitivity of temporal relations, focusing on six types of temporal relations (*simultaneous*, *before*, *immediately before*, *begins*, *ends*, and *includes*) between events as well as between events and time expressions. The participants of TempEval-1 investigated a variety of features ranging from phrase-based syntactic features (Bethard et al. 2007) and dependency parser derived features (Cheng et al., 2007), to knowledge-based features (Puscasu 2007). While most machine learning approaches model temporal relations as local pairwise classifiers ignoring possible inconsistencies among predicted temporal relations, one line of research investigates global optimization of temporal classifiers. (Chambers and Jurafsky 2008a) introduce global constraints over local classifiers using Integer Linear Programming. Similarly, Yoshikawa et al. (2009) propose joint modeling approach using Markov Logic Networks (MLNs). In these works, globally optimized temporal classifiers improved classification accuracies over local classifiers.

Our work on temporal reasoning is most closely related to Yoshikawa et al. (2009) in that MLNs are used for joint modeling of temporal relations. However, our approach is different from their work in two primary respects. First, we introduce new lexical relation features derived from two English lexical ontologies, WordNet (Fellbaum 1998) and VerbOcean (Chklovski and Pantel 2004), as effective predictors of temporal relations between two events. Second, our model addresses a new task introduced in TempEval-2, which is to identify temporal relations between main and syntactically dominated events in the same sentence.

One of the most prevalent problems for NL Visualization is inferring the environment in which the events are taking place. Over the past few years, there have been several projects that address this concern. Zeng et al. (2003; 2005) use a lexical database called a "Descriptionary" to map "visual" words onto objects in a 3D scene. By combining this with frame analysis they were able to produce a virtual scene that was described in a simplified natural language.

The WordsEye system (Coyne et al. 2010, 2001; Sproat 2001) is similar in that it makes use of a large lexical database of descriptive words. In addition, WordsEye supports active verbs by storing a database of armature deformations of a mesh, allowing it to pose a character in a configuration which depicts the event. Inverse kinematics and depiction rules are used to add realism to the scene. Grounding is computed by using concordance lines, extracting verb-object and verb-preposition-object tuples and then ranking through likelihood ratios.

Bilasco et al. (2007) also used a rule-based approach to set a 3D scene. This system provides a framework for adaptations to user interests or to limitations of the device being used. Glass and Bangay (2007; 2008) present an annotation framework for including multimodality and transforming natural language into a 3D virtual environment. By addressing multimodality, the authors also include techniques for aligning audio, including background images, and using constraints to frame the scene.

Another option for automatically adding visualizations to a narrative is automatic illustration. The Story Picturing Engine (Joshi et al. 2006) generates queries based on the sentences by removing stopwords and using a Wordnet similarity metric to select images. Images are filtered and ordered using mutually reinforced rank and an MCMC algorithm to capture some aspects of human behavior. Schwarz et al. (2010) instead use Information Extraction techniques to select keywords for a Flickr query to select images.

## Key Challenges

Narrative visualization for novice writers' text requires devising effective solutions to a range of computational and representational problems: parsing and semantically analyzing ungrammatical text, creating an effective approach to temporal analysis, and developing solutions to the problems of event, dialogue, and frame semantics. We discuss each of these in turn.

### Grammatical Challenges

Narrative text, especially text composed by novice writers, poses significant NLP challenges. First and foremost, a passage is often riddled with grammatical disfluencies, such as real-word spelling errors, missing punctuation, improper or omitted prepositions, incorrect verb tense, and missing auxiliary verbs, among others. This causes the syntactic and semantic parsers, typically trained on reasonably well-formed sentences, to fail frequently. Secondly, narrative text has a very complex rhetorical structure. Since characters in the story often speak to themselves or to other characters, the narrative voice shifts repeatedly during the story. Further, novice writers frequently omit quotation delimiters identifying a change in narrative voice.

Because one of the goals of the NARRATIVE THEATRE is to support creativity, the system needs to be able to make decisions about changes that are made to the story at run time. These differences can also feed into the other components of the system to allow for online processing and reduce computation time for the system.

Basic textual differences, such as those generated by the Unix *diff* command, or Google's diff-match-patch[1] program, don't carry sufficient knowledge for our purposes. For our system, we are mostly uninterested in low level textual differences. Instead, we would like to measure the changes in the narrative structure over time, in order to appropriately contextualize student edits and measure the influence of our visualization.

Since the system is used by novice writers as part of the revision process, there are several challenges to analyzing the structure of the text. As noted above, the grammar and word choices utilized by the student increase the difficulty of processing the text. We also would like to be able to process incomplete text, so that the student can request the visualization at any time during the revision process.

### Temporal Analysis

Narrative text contains descriptions of event sequences that occur along a certain timeline. Correctly recognizing the temporal orders of the events is an essential part of understanding such text. However, human-written text does not always follow a chronological order of events Although events may be expressed by different syntactic categories, such as verbs, nouns, and adjectives in English, the majority of events appear in the form of verbs in the training corpus, and thus the temporal classifier only considers verbs. Given two main verbs in adjacent sentences, the temporal classifier predicts one of six values for their temporal relations: *before*, *after*, *overlap*, *before-or-overlap*, *overlap-or-after*, and *vague*.

### Event Semantics

Event semantics provide the basic unit for both the textual and visual representations of the story. However, different information is conveyed by the same event in the text representation than the visual. In order to align the two representations, we require a translation component that is able to reason about both levels. For this work, we have extended the Narrative Chains model (Chambers & Jurafsky 2008a) by including semantic role information. These semantic frames are then each associated with a *Pose*, a visual configuration of a 3D character depicting the event, in a 1-1 manner.

### Dialogue Semantics

Dialogue is perhaps one of the simplest semantics required for this task. Once dialogue has been identified, any semantic information associated with the spoken text becomes irrelevant. Information within the dialogue, such as anaphoric references, can ground the utterance in the virtual world and aid in identifying the participants of a dialogue. Information within the dialogue, such as anaphoric references, can ground the utterance in the virtual world and aid in identifying the participants of a dialogue. Identification of the dialogue is slightly more challenging than it first appears because both direct

[1] http://code.google.com/p/google-diff-match-patch/

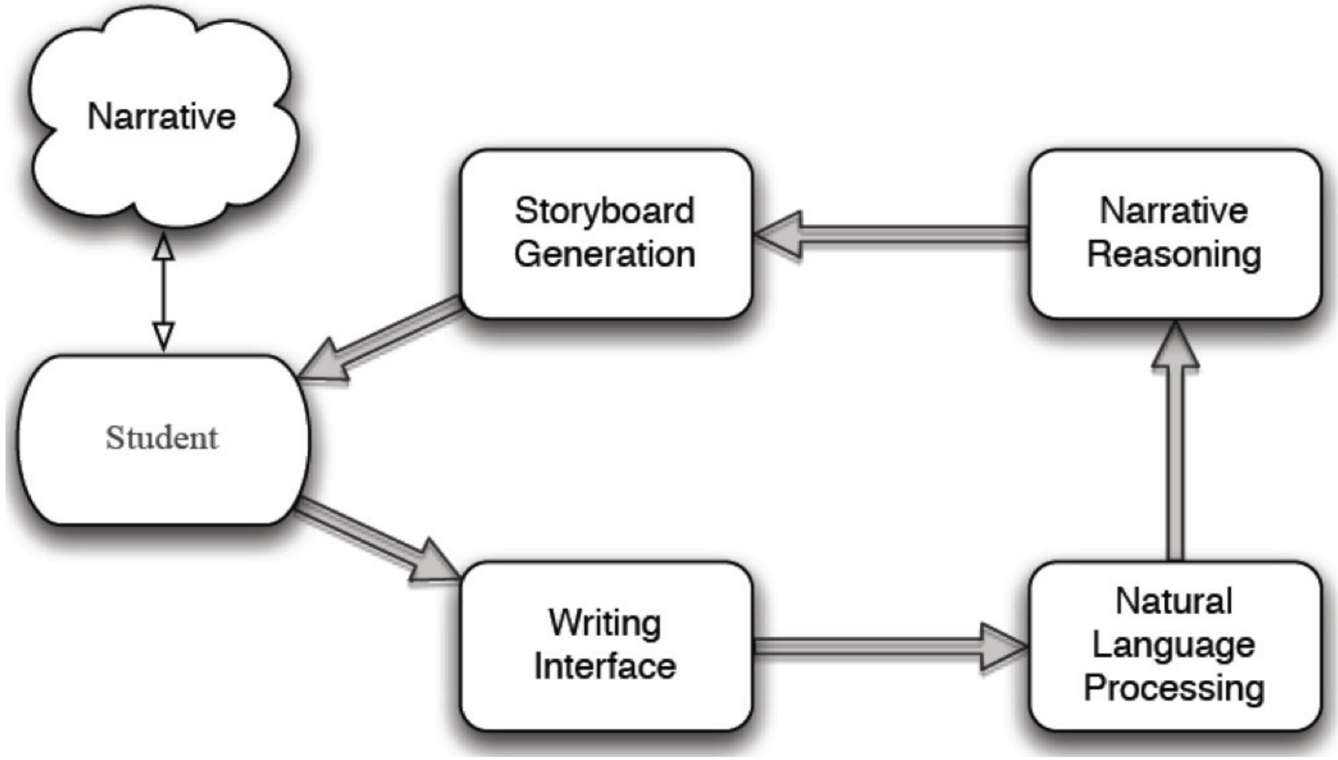


**Figure 1 - NARRATIVE THEATRE architecture**

quotation (“I want to go to the lake,"’ Rabbit said) and indirect quotation (Rabbit said he wanted to go to the lake) can be present in the text. In addition, grammatical challenges as discussed above can hinder the correct identification of a dialogue event.

The other challenge of dialogue in narrative text is attribution. There has been some interest in automatically attributing dialogue to a speaker. Some recent work (Elson and McKeown 2010) found that a machine learning approach involving multiple learners was able to identify and attribute quotes with high accuracy.

### Frame Semantics

The standard frame semantics, such as those that drive planning techniques, are not readily applicable to narrative text. Events can happen “off screen” that change the state of the world beyond what is conveyed in a text event. However, it is also incorrect to ignore frame reasoning. In a fully observable story world, frame semantics are closely related to the notion of cohesion in a story. Through careful omission, a storyteller can generate tension, suspense, or surprise.

In order to properly illustrate the scenes created in the students’ fables, it is necessary to employ some form of coreference resolution in order to determine the character or characters being referred to by certain phrases within the fables. These phrases are most commonly singular and plural pronouns, or names that the students may have given to the characters.

## NARRATIVE THEATRE

We have been exploring these issues in the context of a narrative-centered writing support system for novice writers, the NARRATIVE THEATRE. The NARRATIVE THEATRE is a narrative visualization system designed to support creativity of developing writers. The visual component provides an alternate representation of the story to the students, so that they may be inspired to make changes to their story.

### Narrative Architecture

In the NARRATIVE THEATRE, the computational models dynamically craft visual representations of students’ stories by creating visual storyboards, similar to those used by filmmakers and animators while planning a film. The NARRATIVE THEATRE has four major components: the Writing Interface, the Natural Language Processing (NLP) module, the Narrative Reasoning module, and the Storyboard Generation module.

The Writing Interface is a visual interface designed to lead the writers through the planning, writing, and revision process. First, the writers are asked to select story elements, such as setting and characters, to use in their story. After this selection phase, writers are prompted to plan their stories by making an outline based on the standard three-act Aristotelian plot structure. They are then able to write their fable. During the writing process, all of their previous decisions are viewable from the writing window. They are then presented with the visual storyboard of their work and are allowed to take notes. Finally, they are taken back to the writing screen to do revisions.

In order to support these interactions, the other modules work in conjunction to present the visual storyboard to the writer. The NLP module is responsible for syntactic, semantic and discourse level processing of the students' text. The Narrative Reasoning module is responsible for inferring information not present in the text. This can include frame axioms, consistency reasoning, and the visual semantics reasoning. The visual semantics connect the NLP module with the Storyboard Generation module. It creates a mapping between the event representation induced by the NLP module and the individual storyboard frames for the Storyboard Generation module.

The final module is the Storyboard Generation module. The Storyboard Generation module is responsible for rendering each storyboard frame on screen as a single slide. Each frame contains references to the characters and objects involved in the event, as well as the information needed to visually represent the event. Each event is represented as a single frame of an animation that portrays the character in the middle of performing the event. Each character reference in the event is mapped to a 3D model. Objects and characters are then placed on the screen and the camera is positioned so that all of the characters and objects can be seen, involved in the event can be seen on the screen.

## Corpus Collection

We have collected two corpora of student-written narratives to train the computational models of the NARRATIVE THEATRE. The first focused on myths, while the second focused on fables. Our discussion here focuses on the fable corpus. A multimedia interface was created with Adobe's Flash® development platform and AIR® runtime environment to guide the students through the planning and writing process.

During the planning phase, users select a moral, a setting, a cast of characters, and a set of objects for the story they will create. The system provides nine morals, four settings, ten characters, and twenty objects from which users may choose. Each setting is accompanied with a visual representation. Clicking on an image will open an enlarged image, which is annotated with salient features of the setting. Characters and objects are also visually represented by static graphics. Characters were designed to be neutral in both gender and expression in order to allow users creative choice when filling narrative roles with the characters.

Once the choices have been made, users are presented with a planning area that allows them to view their past decisions and begin structuring their fables. The top of the page contains windows that display the setting, characters, and objects, and that can provide more information on a mouseover by the students. They then craft a textual plan for the beginning (setting and characters are introduced), middle (conflict and problem), and end (conflict resolution) of their stories. After the planning information is entered, the user may begin writing (Figure 2). They then create the actual prose, which is entered as raw text. The writing and revision phases are supported with a spell-correction facility. All student activities including interface selections and the text streams from planning and writing are logged and time-stamped. To avoid the possibility of distracting the student from the writing task, the spell-correction facility was removed in the second version of the system.

The corpus collection activity spanned two days for each student involved. On the first day, the students were seated at a computer and asked to fill out a pre-survey questionnaire, which they were given approximately twenty minutes to complete. On the second day, the students were assigned to a computer again and presented with a login screen for the NARRATIVE THEATRE interface (Figure 2). Once they entered their credentials, the students were presented with a short instructional video that described the features and operation of the interface. They were given fifteen minutes to complete the planning activity described earlier (choosing a setting, main characters, props, and deciding the beginning, middle, and end of their story). Once planning was completed, or time expired, the students were given another thirty-five minutes to write their fables. At the end of this block of time, the students were asked to compete a post-survey questionnaire, for which they were allotted twenty minutes for completion. In total, the session lasted no more than eighty minutes.

## Discussion

Narratives produced by novice writers pose significant NLP challenges because of disfluencies. The stories in the Fable Corpus exhibit significant grammatical problems (Goth et al. 2010). To address this, the NARRATIVE THEATRE architecture includes a grammatical disfluency corrector to mitigate some more common errors. This grammatical disfluency corrector directly addresses punctuation elision as well as real-word spelling errors using a supervised machine learning model. Periods and quotation marks not explicitly expressed, but predicted by

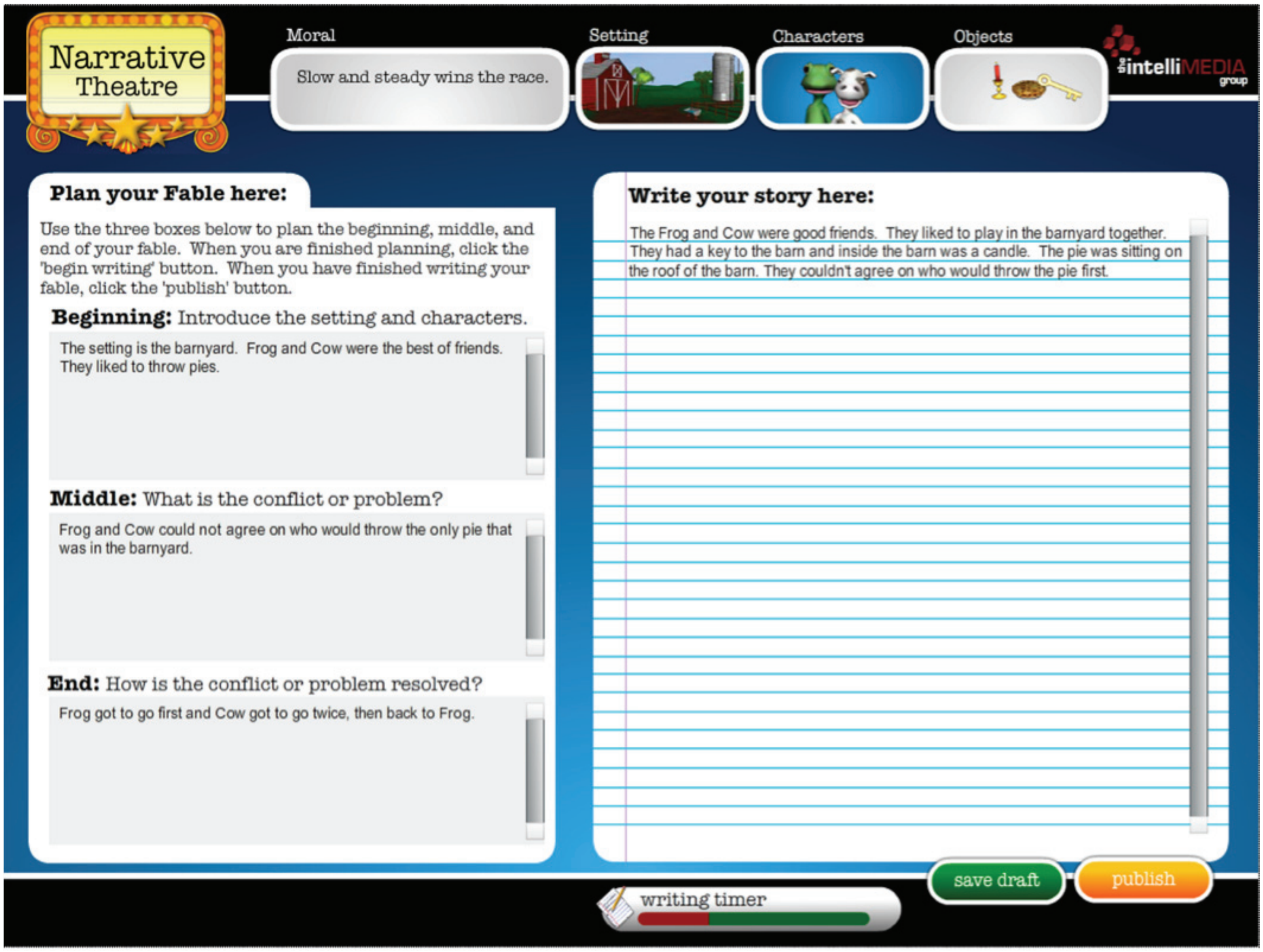


**Figure 2 - NARRATIVE THEATRE interface**

the model to exist, are inserted. Additionally, words predicted to be real-word spelling errors are replaced with the intended word. These corrections are made prior to any syntactic and semantic processing of the student's passage.

As most current natural language tools have been trained on or designed for use with corpora dealing in non-fictional domains, there were some difficulties in using these tools for the domain of fictional fables. These issues stemmed largely from the use of anthropomorphized animals, which served as the main characters of the fables. Tools trained on a corpus of the Wall Street Journal would understandably have few examples on which to base an association of gendered pronouns to foxes, lions, or owls. Similarly, such tools would most likely have no examples of such animals talking, laughing, or playing a drum. In an attempt to alleviate this problem, all instances of the animal characters in the fables were capitalized, so that the tools used might interpret them as proper nouns (and thus as persons), though this approach proved less than fruitful, with only very minor improvements.

Another way we chose to combat the language problems in the corpus was to leverage the strongly constrained domain. Following the middle school language arts curriculum, we focus on fable writing as a constrained writing exercise. By offering the student a selection of expansive but finite sets of morals, characters, and objects, we restrict the possible groundings of events and can reason more easily.

Another benefit of the constrained writing exercise arises in the visualization generation. By limiting the genre and objects, we were able to limit knowledge engineering for the visualization component to events that are intrinsic to the genre or afforded by the objects and characters we presented to the students.

To automatically recognize temporal orders among the events, we constructed a temporal classifier (Ha et al. 2010). The temporal classifier was created using the Markov Logic framework and trained on the *Tempeval-2* corpus (Pustejovsky & Verhagen 2009). The corpus consists of a training set of 162 news articles and is annotated with events, temporal expressions, and temporal relations, as well as manually analyzed semantic attributes (e.g., *polarity*, *modality*, *tense*, *aspect*, and *event class*). Since manually analyzed features are not available from student-written fables, the deployed temporal classifier utilizes only automatically extracted features, such as word tokens and syntactic features.

Events are the core semantic units of a narrative. We considered several options for the event representation. The SCHEHERAZADE system (D. Elson and K. McKeown 2009) was discarded because the annotation time is costly, and the tool developed had low agreement between subjects. We observed that this was a function of the expressivity of this representation. Since there were multiple ways to represent the same narrative sequence, it would be difficult to automatically induce the model.

The Narrative Event Chains model (Chambers and Jurafsky 2009) forms the basis of the event representation we use. The model describes a narrative as a sequence of *Events*, represented by the verbs extracted from the given text. The model, however, is focused on syntactic relationships between the verb and other entities. Since in the NARRATIVE THEATRE we are interested in more closely modeling the semantic relationships between words, we have replaced the syntactic dependency parse information with the more semantically rich semantic role labels. The 1-1 mapping of semantic frames to events is overly simplified, as in some cases, the visual representation has the ability to capture multiple textual events in a single visual frame. In others, multiple visual frames might be required to convey a single textual frame. Exploring such connections is an area of future work.

The NARRATIVE THEATRE system will also benefit from the development of a question generation component, providing the student with interactive feedback on their writing performance. Questions can range in specificity, from inquiring on disfluencies in grammatical construction, to deeper Socratic categories of questions.

## Conclusion

Narrative visualization is the task of dynamically creating visualization of narratives. It poses significant computational challenges for well-written narratives and is even more challenging for narratives composed by novice writers. We have presented an emerging framework for narrative visualization and introduced a narrative environment, the NARRATIVE THEATRE, that is being constructed to explore narrative visualization. It is being designed as a writing support environment for middle-school writers for the domain of fables. We are actively exploring approaches for each of the primary computational problems posed by narrative visualization, and plans are underway to empirically explore how the narrative visualization can support novice writers.

## Acknowledgements

The authors wish to thank members of the North Carolina State University IntelliMedia Group for their assistance with NARRATIVE THEATRE. This research was supported by the National Science Foundation under Grant IIS-0757535. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

## References

Allen, J.F. An Interval-Based Representation of Temporal Knowledge. *Proc. 7th International Joint Conference on Artificial Intelligence, Vancouver, Canada*, 1981, 221–226.

Bethard, S., Martin, J.H., and Klingenstein, S. Timelines from Text: Identification of Syntactic Temporal Relations. 2007.

Boguraev, B. and Ando, R.K. TimeML-Compliant Text Analysis for Temporal Reasoning. *Proceedings of IJCAI*, 2005, 997–1003.

Chambers, N. and Jurafsky, D. Jointly Combining Implicit Constraints Improves Temporal Ordering. *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, Association for Computational Linguistics 2008a, 698–706.

Chambers, N. and Jurafsky, D. Unsupervised Learning of Narrative Event Chains. *Proceedings of ACL-08: HLT*, 2008b, 789–797.

Chambers, N. and Jurafsky, D. Unsupervised Learning of Narrative Schemas and their Participants. *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 2*, Association for Computational Linguistics 2009, 602–610.

Chklovski, T. and Pantel, P. Verbocean: Mining the Web for Fine-grained Semantic Verb Relations. *Proceedings of EMNLP*, 2004, 33–40.

Coyne, B., Rambow, O., Hirschberg, J., and Sproat, R. Frame Semantics in Text-to-Scene Generation. *Knowledge-Based and Intelligent Information and Engineering Systems*, 2010, 375–384.

Coyne, B., Sproat, R., Dupuy, S., Egges, A., Legendre, V., and Nugues, P. WordsEye: An Automatic Text-to-Scene Conversion System, SIGGRAPH 2001. *Proceedings of the 28th Annual Conference on Computer Graphics, Los Angeles, California, USA. ACM*, 2001.

Elson, D. and McKeown, K. Extending and Evaluating a Platform for Story Understanding. *Proceedings of the AAAI 2009 Spring Symposium on Intelligent Narrative Technologies II*, 2009.

Elson, D.K. and McKeown, K.R. A Platform for Symbolically Encoding Human Narratives. *Proc. of the AAAI Fall Symposium on Intelligent Narrative Technologies*, 2007.

Fellbaum, C. *WordNet: An electronic lexical database*. The MIT press, 1998.

Glass, K. and Bangay, S. Automating the Creation of 3D Animation from Annotated Fiction Text. *Proceedings of the IADIS International Conference on Computer Graphics and Visualization*, 2008, 3.

Goth, J., Baikadi, A., Ha, E., Rowe, J., Mott, B., and Lester, J. Exploring Individual Differences in Student Writing with a Narrative Composition Support Environment. *Proceedings of the NAACL HLT 2010 Workshop on Computational Linguistics and Writing: Writing Processes and Authoring Aids*, Association for Computational Linguistics 2010, 56–64.

Ha, E.Y., Baikadi, A., Licata, C.J., Mott, B.W., and Lester, J.C. Exploring the Effectiveness of Lexical Ontologies for Modeling Temporal Relations with Markov Logic Modeling Temporal Relations with Markov Logic. *Eleventh Annual SIGDIAL Meeting on Discourse and Dialogue*, 2010, 75-78.

Johansson, R. and Berglund, A. Automatic Text-to-Scene Conversion in the Traffic Accident Domain. *Proceedings of the Nineteenth International Joint Conference on Artificial Intelligence*, 2005, 1073-1078.

Joshi, D., Wang, J.Z., and Li, J. The Story Picturing Engine—a System For Automatic Text Illustration. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP) 2*, 1 2006, 68–89.

Puscasu, G. Wvali: Temporal Relation Identification by Syntactico-Semantic Analysis. *Proceedings of the 4th International Workshop on SemEval*, 2007, 484–487.

Pustejovsky, J., Castano, J., Ingria, R., Sauri, R., Gaizauskas, R., Setzer, A., Katz, G., and Radev, D. TimeML: Robust Specification of Event and Temporal Expressions in Text. *IWCS-5 Fifth International Workshop on Computational Semantics*, 2003.

Schwarz, K., Rojtberg, P., Caspar, J., Gurevych, I., Goesele, M., and Lensch, H. Text-to-Video: Story Illustration from Online Photo Collections. *Knowledge-Based and Intelligent Information and Engineering Systems*, 2010, 402–409.

Sproat, R. Inferring the Environment in a Text-to-Scene Conversion System. *Proceedings of the 1st international conference on Knowledge capture*, ACM 2001, 147–154.

Ye, P. and Baldwin, T. Towards Automatic Animated Storyboarding. *Proceedings of the 23rd national conference on Artificial intelligence*, 2008, 578–583.

Zeng, X., Mehdi, Q., and Gough, N. From Visual Semantic Parameterization to Graphic Visualization. *Information Visualisation, 2005. Proceedings. Ninth International Conference on*, IEEE 2005, 488–493.