

Building a Bank of Semantically Encoded Narratives

David K. Elson and Kathleen R. McKeown

Columbia University, New York City
{delson, kathy}@cs.columbia.edu

Abstract

We propose a methodology for a novel type of discourse annotation whose model is tuned to the analysis of a text as narrative. This is intended to be the basis of a “story bank” resource that would facilitate the automatic analysis of narrative structure and content. The methodology calls for annotators to construct propositions that approximate a reference text, by selecting predicates and arguments from among controlled vocabularies drawn from resources such as WordNet and VerbNet. Annotators then integrate the propositions into a conceptual graph that maps out the entire discourse; the edges represent temporal, causal and other relationships at the level of story content. Because annotators must identify the recurring objects and themes that appear in the text, they also perform coreference resolution and word sense disambiguation as they encode propositions. We describe a collection experiment and a method for determining inter-annotator agreement when multiple annotators encode the same short story. Finally, we describe ongoing work toward extending the method to integrate the annotator’s interpretations of character agency (the goals, plans and beliefs that are relevant, yet not explicitly stated in the text).

1. Introduction

This paper describes a methodology for constructing a novel type of language resource consisting of a corpus of annotated textual stories. The annotation takes the form of a single interconnected graph that encodes the entire discourse. Nodes in the graph represent spans of the reference text, objects in the story, and propositions (predicate-argument structures) that convey actions, statives, and modifiers. All of the symbols are linked to formal ontologies. Nodes are connected by arcs that represent textual connections such as causality and motivation. In addition, we describe an initial collection experiment and a method for measuring inter-annotator agreement. We detail in (Elson and McKeown, 2009) an annotation tool called SCHEHERAZADE that elicits such *story graphs* from non-expert annotators. In this paper, we focus on the procedure for encoding a story graph from a reference text.

We have been motivated to develop this annotation scheme by the lack of an existing language resource that covers all these facets of a text. For example, large-scale annotation projects have focused on identifying predicates and arguments at the sentence level (Kingsbury and Palmer, 2002) or the connections between sentences at the discourse level (Prasad et al., 2008; Carlson et al., 2001). Temporal and modal properties of a corpus have been separately modeled (Pustejovsky et al., 2003). We are interested in a more holistic view of a story, so that the interplay of many features can be understood – our model involves word sense disambiguation, coreference, interpretation of tense and aspect into a symbolic view of time and modality, and the encoding of actions, statives and modifiers into propositional structures. In addition, we are currently extending the model to include crucial elements of character agency (goals, plans and beliefs) even when they are not stated in the text. We believe that a corpus of narratives annotated into story graphs (a “story bank”) would galvanize work in this area of text understanding, and allow us to develop the tools that reveal how stories work: the structure that the narrator chooses, the roles that characters take on, the similarities between stories and across genres, and so on.

In the following sections, we describe our method for synthesizing several modes of annotation into a single procedure in which annotators build a structure that reflects the content of a reference text. Section 2 reviews our formal representation. We outline the three stages of annotation in Section 3: extraction of objects and themes, construction of predicate-argument structures, and assembly into a single interconnected graph. In Section 4, we describe a collection experiment and a method for calculating inter-annotator agreement that measures the similarity between two story graphs. Section 5 describes ongoing work in eliciting implied information about character agency, which was not included in the collection experiment. Finally, we review related work and conclude.

2. Story graph approach

Following theorists such as (Bal, 1997), we define a “narrative” discourse as one which refers to events that transpire in the world of the story, ordered with temporal relations and linked by threads of causality and motivation. Our methodology guides annotators toward encoding the underlying events present in a story, rather than the lexical choice or other elements of the story’s realization in text. In an earlier paper (Elson and McKeown, 2007), we outline our method of encoding these events with a combination of propositions (predicate-argument structures) and conceptual graphs. While this paper focuses on the encoding process from the annotator’s perspective, this section summarizes our approach to story representation.

Nodes in the conceptual graph represent actions, statives, modifiers, time states, spans of reference text and metadata such as the characters, locations and props found in the text. Action, stative and modifier nodes contain definitional propositions. The predicate-argument structures are adapted from the VerbNet lexical database (Kipper et al., 2006) while a taxonomy of nouns and adjectival modifiers are supplied by WordNet (Fellbaum, 1998). The arcs of the conceptual graph include temporal relationships (which actions take place at which time states or intervals), direct causal relationships (which travel from an action to its

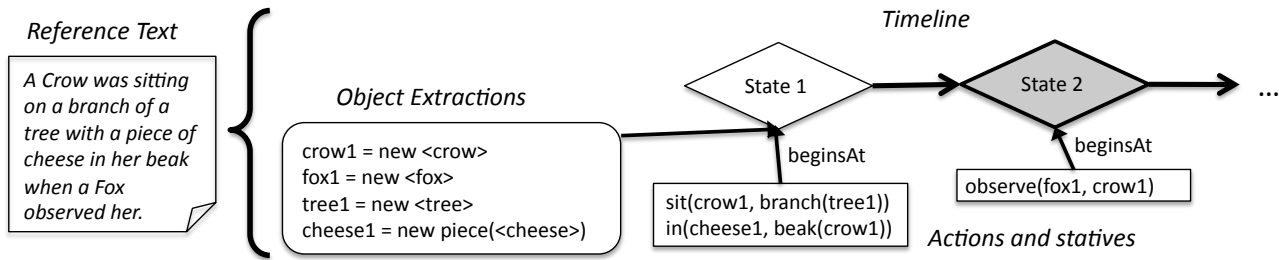


Figure 1: Schematic of a portion of a story graph.

direct consequence actions, if any) and equivalence relations between spans of reference text and their corresponding propositions. Not all propositions need to describe actual story events; actions, statives and even entire “alternate timelines” can be set to modalities including the obligations, beliefs and fears of characters. For example, a character can hope that something will not happen in the future or has not happened in the past. Figure 1 illustrates a portion of a story graph schematic, showing the first sentence of a reference text, as well as objects, statives, actions and points in time identified in the text by an annotator.

The reference text we will use in this paper is one of the fables attributed to Aesop. We have developed our methodology using Aesop’s fables as a development corpus, but our knowledge sources are not tuned to these stories or their domain. In this paper, we will discuss a fable titled *The Fox and the Crow*, which has been featured in other story representation schemes (Ryan, 1991), and is reproduced below:

A Crow was sitting on a branch of a tree with a piece of cheese in her beak when a Fox observed her and set his wits to work to discover some way of getting the cheese. Coming and standing under the tree he looked up and said, “What a noble bird I see above me! Her beauty is without equal, the hue of her plumage exquisite. If only her voice is as sweet as her looks are fair, she ought without doubt to be Queen of the Birds.” The Crow was hugely flattered by this, and just to show the Fox that she could sing she gave a loud caw. Down came the cheese, of course, and the Fox, snatching it up, said, “You have a voice, madam, I see: what you want is wits.”

One important aspect of the story graph is that it is a static model of the entire story, without a “beginning” or “end.” The telling of the story – that is, the ordering of propositions that constitutes a *narration* of the graph from a beginning to an ending – is given through the arcs that the annotator provides between the spans of reference text and their related proposition(s). As one progresses through the reference text in a story graph from start to finish, one will encounter arcs to propositional nodes attached to the symbolic timeline. Many narratives, including Aesop’s, reveal the actions in their timelines in a nonlinear order (i.e., with jumps to the past or to the future and back again). While this fable is strictly linear, the fox’s dialogue refers to a hypothetical future.

2.1. Encoding tool

We have implemented and publicly released a graphical annotation tool that facilitates the process of encoding a reference text into a story graph. Figure 2 shows the main screen as configured for annotating *The Fox and the Crow*. The methodology we describe below takes place in the context of the buttons and menus offered by this tool. The central panel shows a timeline with states (called “Story Points”) arranged in a linear graph. Alternate timelines can also be accessed from this screen. The top-right and bottom-right panels contain the reference text and a *feedback text*, respectively. The latter is an automatically generated realization of the story graph that uses syntactic frames from VerbNet and a model of tense and aspect assignment. It facilitates the encoding process by giving annotators a natural-language equivalent of the story graph which they can compare to the reference text at every step.

3. Annotation Procedure

The first task for an annotator is to read the text and fully comprehend it. As a story graph is a discourse unit rather than a sentential unit, it is important for the annotator to take a holistic view of the text before beginning the encoding process.

The method for creating a story graph from a reference text involves three tasks. They need not occur sequentially; an annotator can move back and forth as needed. After making each change to the story graph, the annotator checks the feedback text to ensure that the system has correctly encoded the concept. The process terminates once the annotator feels that he or she has encoded the story with the greatest amount of precision possible, given the formal limitations of the representation. The tasks are:

1. **Object extraction.** Named entities, objects and relevant themes or processes in the text are identified, as well as their immutable statives (inherent properties).
2. **Construction of propositions.** Predicates are selected for each action and stative. Arguments are then filled in for each thematic role associated with each predicate. Modifiers and negations, if any, are also assigned.
3. **Assignment and linking.** Propositions are assigned to states and transitions on a timeline and linked to corresponding spans of reference text. Modal (hypothetical) events are grouped in separate timelines.

In this section, we describe each task in detail.

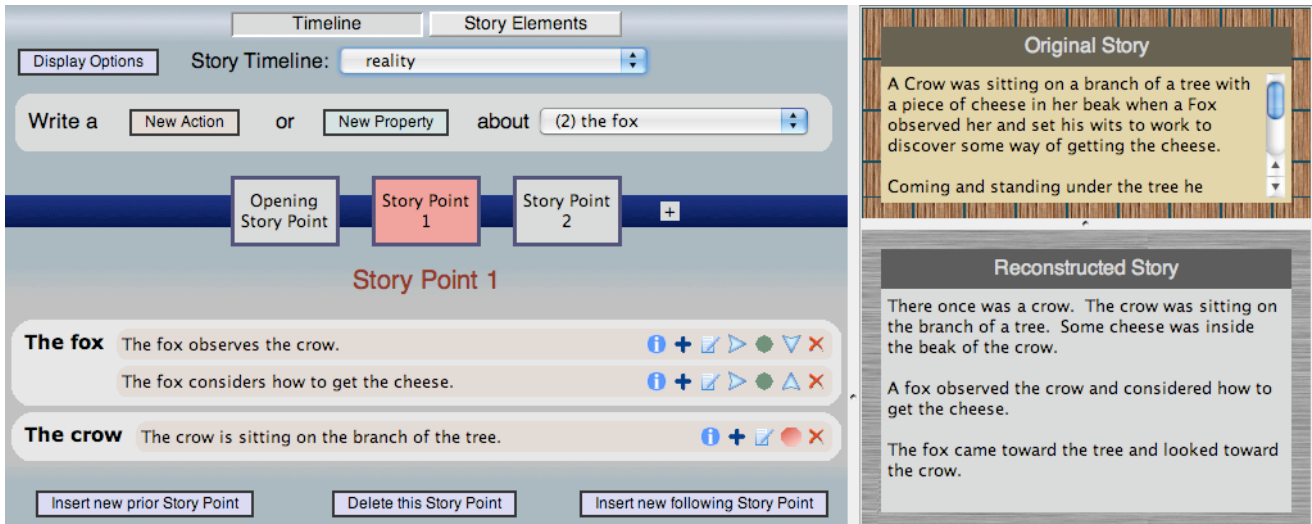


Figure 2: Main screen of the SCHEHERAZADE interface.

3.1. Object and theme extraction

After reading the text, the annotator first identifies the “named entities” that occur. We use this term to refer to objects and themes that are named by the annotator – even if the story’s narrator does not assign names – so that they can be invoked in multiple propositions. In other words, the annotator performs coreference on the text by identifying entities and assigning them unique identifiers; each mention of the entity is then invoked by the identifier and linked to the other mentions.

Figures 1 and 2 indicate the entities that are appropriate for an annotator to extract from *The Fox and the Crow*: a crow, a fox, some cheese, and a tree. The annotation tool allows annotators to search for such entities from among a list populated by a subset of WordNet (those synsets which can serve as characters, locations, props, processes, and so on). Figure 1 gives each instantiated object an identifier such as *crow1*, but unless the annotator provides a name (such as *John*), these are not exposed in the graphical interface. Whenever a predicate calls for an *Actor* as an argument, for instance, the interface presents a menu of all animate characters that the annotator had extracted (such as in Figure 5). The feedback text generator selects the correct determiner and uses numerical disambiguators (*the first man*, *the second man*) to reflect the coreference modeled in the story graph. For example, in this story, all mentions of a fox should refer to the same fox entity; only the first should be realized as *a fox*, signaling a new character.

The fable in question includes several objects and themes which only occur once. The annotator can declare a named entity for these in the same manner, or simply invoke them “on the fly” while constructing the propositions in which they appear. In our reference fable, such cases include the branch of the tree (which is modeled in Figure 2), the plumage of the bird, the hue of that plumage, and intelligence (in the abstract). Objects and themes can be used in composition, with an “of” relationship inferred: *hue(plumage(bird))* refers to the hue of the bird’s plumage. Previous work in word sense disambiguation has shown WordNet synsets to be very fine-grained, which can hurt

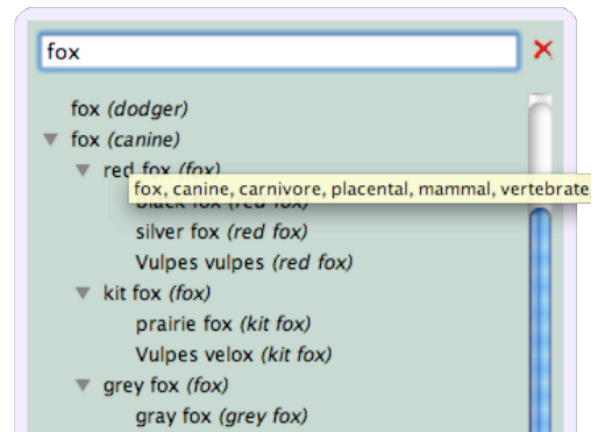


Figure 3: Searching in the character extraction panel.

inter-annotator agreement (Palmer et al., 2007). For example, “fox” also represents the synset of “a shifty deceptive person”; other distinctions can be far more subtle. We address this issue in two ways. First, at the user interface level, we make it clear when there is an ambiguity, and trace the hyponymy of each synset. The annotator sees “fox, canine, carnivore” juxtaposed to “fox, deceiver, wrongdoer.” Figure 3 shows a hyponymy trace appearing when the annotator hovers the cursor over a particular result when searching for “fox” during character extraction. The second technique is to limit ambiguities by selectively disassociating certain words with their less-used synsets. Specifically, we set a threshold for the information content a synset must have to serve as an option for a word. This is calculated from a model of the usage of each synset in a naturally occurring corpus.

Annotators may invoke abstract entities. For instance, a character may refer to *a tree* rather than some particular tree, named or otherwise. They may also refer to groups of entities in the singular. In *The Fox and the Crow*, this is necessary several times: first, when the fox refers to the crow as *a noble bird*, and second, to a hypothetical *Queen of [all] the birds*.

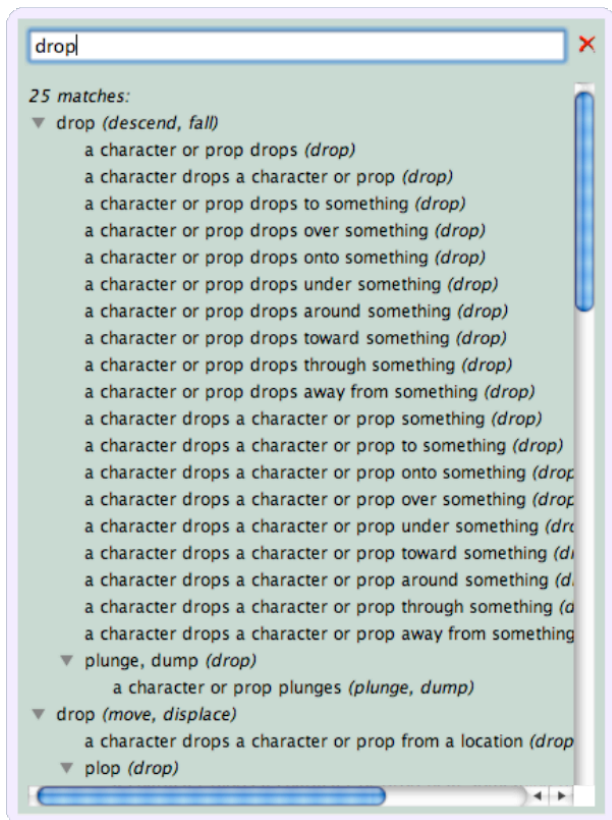


Figure 4: Predicate selection panel showing search results for “drop.”

3.2. Predicate and argument selection

After the annotator extracts named entities, she chooses the best predicates and arguments to describe the actions, stative and modifiers that transpire in the text.

VerbNet has also been mapped to the large-scale annotation project PropBank (Kingsbury and Palmer, 2002). There are three key differences between PropBank and this project. First, PropBank focuses on the sentence level; the result of a story encoding is a single structure consisting of many interconnected propositions. Second, there is not necessarily a direct mapping from predicates in the reference text to propositions in the story graph. As we will see, annotators may alter or consolidate text to focus on the underlying story events rather than the rhetoric. Third, PropBank fills its arguments with spans of text, such as assigning a clause to Arg0 or Arg1. Our representation discourages this use; normally, an annotator chooses an element from a formal taxonomy (such as a named entity) to serve as an argument. For example, rather than highlight the text *the window*, the annotator would select a named entity declared from the correct WordNet *window* synset. If no formal symbol can be found for an argument, the annotator can indicate a text span as a last resort.

At the user interface level, the annotator searches for a verb, stative or modifier predicate and is then presented with a list of frames that match the query. The generation component realizes an example phrase using placeholder arguments as an aid for the annotator. For example, a search for “drop” offers several frames with varying thematic roles: *A char-*

acter or prop drops, A character or prop drops a character or prop, and others (see Figure 4).

Once the annotator selects the predicate, the interface prompts her to select arguments. The selection panel changes depending on the thematic role of the argument: an *Experiencer* argument would present a list of extracted characters, while a *Stimulus* argument which VerbNet reports as compatible with a *Communication* would prompt the user for a predicate to serve as the basis for a nested dialogue proposition. Figure 5 shows the prompt for the dropped object invoked by a “drop” predicate.

Sometimes the action or stative choice is clear and obvious. For example, the reference text includes the clause *when a Fox observed her*. To encode this, the annotator searches for “observe” and finds “something observes something.” She can then select the observing thing and the observed thing. The feedback text replies, *The fox observed the crow*.

At other times, a fair amount of simplification is necessary. For example, our guidelines have annotators phrase passive voice statements as active voice where possible. In *The crow was hugely flattered by this*, the annotator would model it as if it said *The fox flattered the crow*. While the fox’s statements are what truly flattered the crow, we ask annotators to simplify in cases such as these, creating a “canonical” form that facilitates later analysis.

However, there is a qualitative difference introduced in this rephrasing. In *The crow was hugely flattered by this*, it is clear that the flattery is the crow’s belief. *The fox flattered the crow* reads as though the fox’s intention was to flatter the crow and he succeeded in doing so. The correct meaning can be made clear by making the proposition an argument to an encapsulating belief proposition: *The crow believed that the fox had flattered it*.

There are other examples in this fable where forcing the annotator to choose arguments from a controlled vocabulary steers her to focus on content rather than rhetoric. Idioms and figures of speech are replaced by simpler synonyms. For instance, the story includes the following:

Coming and standing under the tree he looked up and said, “What a noble bird I see above me!”

An appropriate encoding of this sentence would be:

```
standUnder(fox1, tree1)
say(fox1, crow1, noble(crow1))
```

These propositions include only the essential details of the sentence. It is not important that the fox identifies the crow as a crow (that is never in question), just that the fox declares the crow to be noble. The wording of the fox’s dialogue is stripped of its flourish. (While “stand under” is not in WordNet, “under” is one of the prepositions listed as being compatible with “stand” in VerbNet.)

We do not prescribe some number of propositions for the annotator to construct. Rather, the annotator uses as many as needed to fully capture the story content (to the extent permitted by the formal restrictions). A sentence may map to more than one proposition.

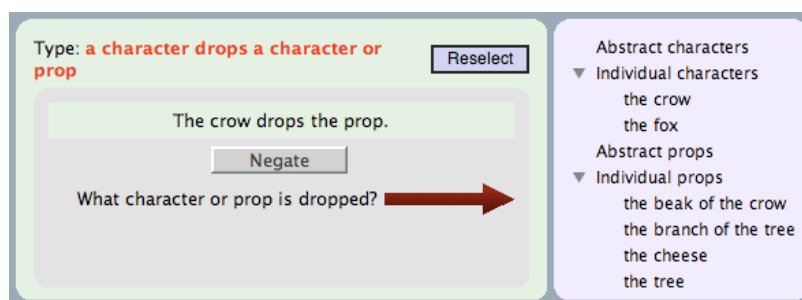


Figure 5: An argument selection panel, which is shown once a predicate is selected.

3.3. Time, mode, connectives

Once the annotator constructs individual propositions, she must attach them to the overall scaffold of the story graph. The connective properties we ask annotators to consider include the following:

1. **Temporal** relationships. Each action and stative is linked to one of the discrete states that make up the main timeline. Actions may be instantaneous or continuous over a span of time. Two propositions share a state or interval only if they occur simultaneously. The timeline represents the order of events as they occur in the story-world, as opposed to the order in which events are told in the original discourse. Annotators may assign text spans to be unbounded, with no beginning or no ending. We plan to add support for habitual and cyclical events in the near future.
2. **Modal** labels. An action and stative can be marked as being *hypothetical* rather than actual, especially if they are conditionals or occur as an obligation, goal or belief of a character (Ryan, 1991). For example, in the familiar fable of the *Milkmaid and Her Pail*, the title character imagines an elaborate future for herself (counting her eggs before they are hatched). Sequences of hypothetical actions can be collocated in a single imaginary timeline.
3. **Modification**. Modifier propositions are linked to the actions, statives or other modifiers to which they apply.
4. **Causality**. If the text makes it clear that one action happens as a direct consequence of another, such as with the “so” and “because” key words, the annotator marks a causal edge from the cause to the result.
5. **Reference text spans**. The entire reference text is included in the story graph, and spans of text can be connected with arcs to the nodes that approximate them. The correlation of a reference text to its encoded equivalent allows us to combine lexical and syntactic features with proposition-oriented features when processing story graphs. As we mentioned earlier, the links to the text spans also preserve the “telling” of the story, in terms of the order in which propositions are narrated – a crucial variable for reader response.

In the case of *The Fox and the Crow*, the temporal flow is monotonic; no flashbacks or flash-forwards are involved.

Each event is assumed to occur subsequent to the one before, except when the word *and* is used in the text, in which case two propositions are assumed to co-occur. The use of *and* is a shorthand, however – it does not prove or disprove a temporal relationship. The annotator must decide the most likely span for each event and determine when events co-occur based on all available evidence. An alternate timeline can be used to indicate that an event took place at some point in the “past” or “future” whose exact placement with respect to the other events in the story is unknowable.

The most significant causal link in this fable runs from the singing of the crow to the falling of the cheese. Our guidelines advise annotators to indicate a causal link if the text explicitly mentions such a connection, or if one event is the sole reason that another occurs at all (or the joint reason with other propositions). Adjacent segments of dialogue are not necessarily causally connected, but a character’s action can be a consequence of another character’s speech.

Figure 6 shows the result of these three steps on *The Fox and the Crow*. The two columns of nodes represent reference text spans and action/stative nodes; the arrows that cross from one column to the other indicate equivalence. The vertical arrows and braces among the left sides of the propositions show the temporal arrangement; for simplicity, all actions and statives are illustrated as sequential or simultaneous points rather than overlapping intervals. The remaining arrows are causal edges.

4. Measuring agreement

In this section, we describe an initial collection experiment and a measurement of inter-annotator agreement that depends upon a graph similarity metric. The notion of inter-annotator agreement is complex in the case of the story graph, for several reasons. First, the graph itself contains a multitude of selections from among very large controlled vocabularies. Second, there is a significant amount of subjective interpretation built into a story graph. Although the guidelines strive to make clear rules for “cutting through rhetoric,” a story is a *received* construction, and varying inferences are to be expected. The English language also permits cases of total ambiguity for which a judgment call is necessary. This dynamic lends itself to a study of varying subjective interpretations among readers; however, in the present case our aim is to find the commonalities between two parallel story graphs of the same reference text.

Our initial collection project involved two encoders who

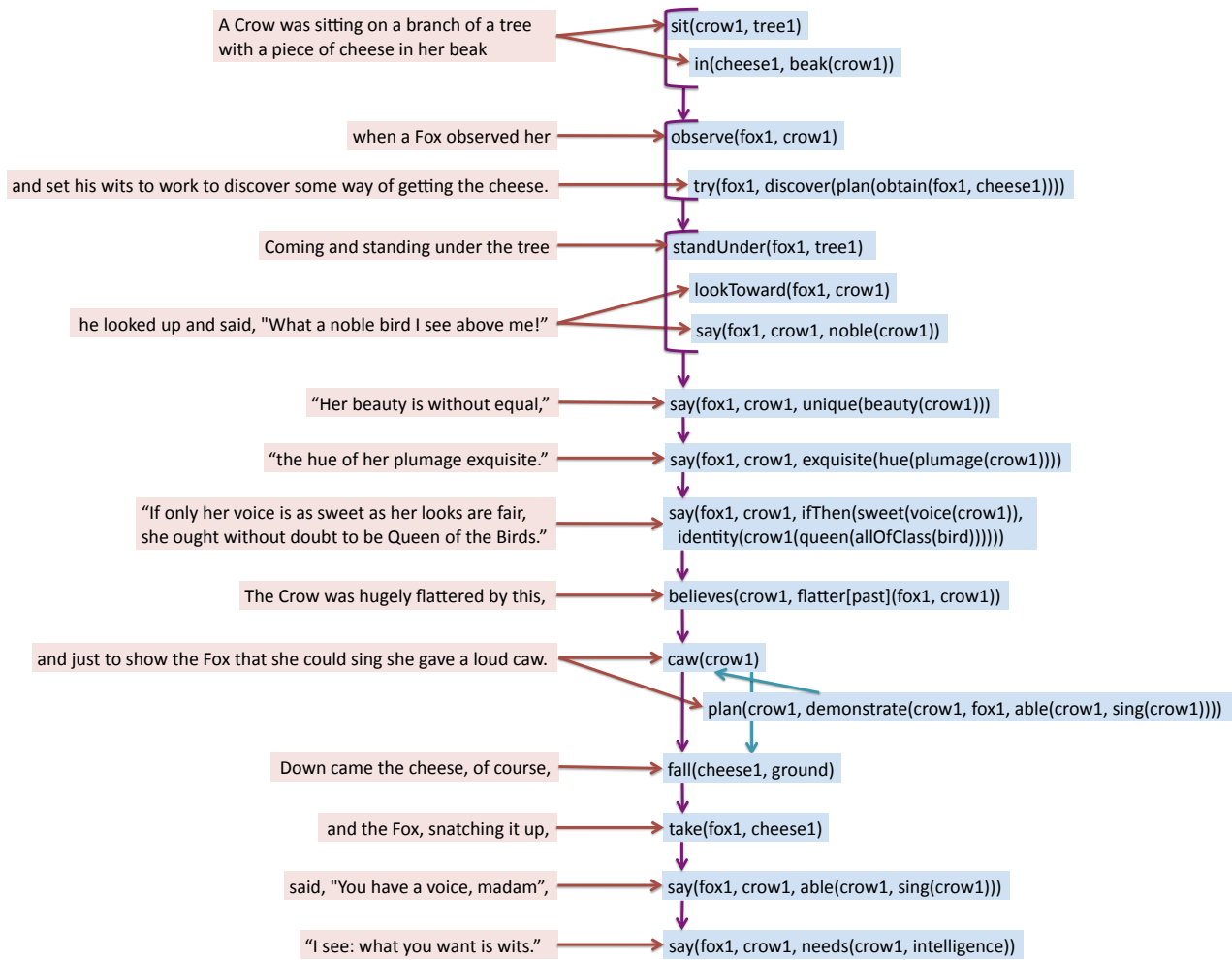


Figure 6: Complete encoding for *The Fox and the Crow*.

each created story graphs for a set of 20 of Aesop’s fables. Both subjects were undergraduates and native English speakers. It took several hours to guide them through training and an initial encoding; as they became familiar with the tool, the required time per fable (80 to 175 words) dropped to 30-45 minutes. The annotators created an average of one proposition for every 9 words of reference text. To measure inter-annotator agreement, we begin with a metric for comparing two propositions or objects. We adopt Lin’s (1998) information-content metric for measuring the semantic similarity of two WordNet senses (those of the predicates or object types, respectively). For propositions a and b , this measure constitutes half of the similarity score $s(a, b)$; the balance is the similarity between the respective arguments, which are recursively scored in the same manner:

$$s(a, b) = \frac{c(a, b) + \frac{\sum_{i=1}^{r(a,b)} s(p(a,b,i))}{r(a,b)}}{2} \quad (1)$$

where $c(a, b)$ is information-content similarity between the two predicates; if a and b are propositions, $r(a, b)$ is the number of arguments that each contains, and $p(a, b, i)$ retrieves the two arguments in a and b for some thematic role i (e.g., the two *Agent* arguments for some two actions). If a and b are objects, the recursive term is ignored.

We then compute the overall similarity between two story graphs by algorithmically aligning the two graphs together, matching the analogous propositions (that is, the pairs that most likely encode the same concept from the reference text). After finding the average similarity between each matched pair, we assign a score penalty for the “singleton” propositions in either graph for which analogues were never found.

To calculate the kappa coefficient (Cohen, 1960) for measuring similarity above chance agreement, we must first determine chance in this context. We do this by generating two *random* story graphs – picking both predicates and arguments at random from among the lexemes found in the corpus – and finding their similarity to one another using the same method as described above.

The resulting kappa is .51, which indicates moderate agreement. We believe this is an encouraging sign that annotators can reliably extract propositions from a text, even when the concepts are presented rhetorically. We further believe that we can improve kappa by implementing heuristics for detecting and normalizing certain types of “semantic paraphrases,” when slightly different predicates are used to encode the same concept from the reference text. Our similarity algorithm already connects verb actions and their derived adjectival statives.

5. Interpretative assessments

All the annotations we have discussed have been based on story content which is stated in the discourse itself. In this section, we outline ongoing work in extending our model and methodology to capture the interpretative connections that give a story its cohesion.

The factors that give *The Fox and the Crow* its moral “point” are left unstated: that the crow’s vanity was its downfall, and that the fox’s ultimately successful plan was always to capitalize on this character flaw. More generally, a skilled narrator does not always state each character’s goals and motivations, but these are essential aspects of a story as it is understood by a reader (Bruner, 1991). Many of the interesting connections between stories do not involve similarities between actions, but rather those between the goals, the plans and the conflicts that arise.

We are currently extending our methodology to capture three key aspects of character agency: goals, plans and beliefs. We call this the *interpretative layer* of the story graph, as opposed to the *stated layer*. Our purpose is to understand the motivation, if any, behind each stated proposition. We draw from theorists since Aristotle who have noted that dramatic narratives hinge on the disconnects between what characters know and how they act. Annotator interpretations, of course, will vary more as we depart from the written word. A story graph becomes more of a model of the received story than an absolute record of story content.

Structurally, the interpretive layer consists of more propositions attached to timelines. The difference is that each is marked as being part of a goal, plan or belief that is unstated and relevant to the story. A goal is a truth value that a character desires to come to pass; a plan indicates a sequence of (possibly branching) actions intended to reach a goal; a belief illustrates the character’s implied world view. These nodes are connected to nodes in the stated layer of the graph. The arcs that span the two layers include:

1. **Attempt to cause:** Indicates that the motivation behind an action is to cause a goal or plan to occur.
2. **Precondition for:** Sets one aspect of a plan as a precondition for the transpiring of a goal or another plan.
3. **Would cause:** Indicates belief that one proposition, should it come to pass, would cause another.
4. **Fulfills:** Links an action to a goal or plan that it actualizes.

The methodology for assigning interpretative-layer nodes is to specify the most relevant goals, plans and beliefs that act as motivations or explanations for the stated content. *The Fox and the Crow* is notable for its intersecting plans. Each character has a motive for its actions; the key to the story is that the crow’s plan to impress the fox is itself part of the fox’s plan to get the cheese. The fox does not state his plan at the beginning of the story; the “unfolding,” where the fox’s plan is revealed upon its actualization, happens at the end. The middle of the story consists of the fox taking actions to actualize an unstated plan.

We have run a formative evaluation in eliciting the interpretative layer of the story graph, involving experts in literary

theory. We found that the expressiveness of the model was sufficient for capturing motivations behind actions in five stories in the Aesop corpus, and that different annotators identified disjoint but similar motivations.

6. Related work

SCHEHERAZADE functions as a WYSIWYM (what you see is what you meant) tool, in that it allows domain experts rather than knowledge engineers to work in a formal representation – with the assistance of a graphical interface and real-time, automatic feedback text (Power et al., 1998). Biller et al. (2005) describe a WYSIWYM editor in which the representation, like ours, is a conceptual graph that incorporates the VerbNet and WordNet linguistic resources. However, compared to this work, our work is geared toward narrative, and includes more complete representations of time, modality and agency.

Conceptual graphs have previously been considered as a formalism for interpreting story content, both in theory (Bearman and Stovel, 2000) and practice (Min et al., 2008). The QUEST approach, in particular, models a story in terms of action/stative nodes and arcs that represent relations such as causality (Graesser et al., 1991; Christian and Young, 2003). However, we know of no previous collection project to allow annotators to encode stories in the QUEST formalism; while the recent ReQUEST project uses the technique for computer/human co-authoring (Riedl et al., 2008), the use of a WYSIWYM editor in this context is novel.

The conceptual graph is by no means the only representation proposed for modeling narratives; to name a few, planning has been applied to plot generation (Riedl and Young, 2005), and logic-based formalisms can capture aspects of commonsense reasoning (Mueller, 2004), psychology (Hobbs and Gordon, 2005) and character agency (Rapaport, 1986). These formalisms are more suited for automatic computation of story content, and can be rigid in their coverage, where our notion of a story graph is more descriptive in nature.

Inference of agency is an active area in its own right (Pollack, 1990). While computational analysis of style is a useful tool for literary scholarship (Mostellar and Wallace, 1984; Lee, 2007), our intention is to assist in the analysis of deeper structural and thematic connections within and between narratives.

7. Conclusion and Future Directions

In this paper, we described a methodology for collecting detailed formal encodings of narratives. The model exposes a feature space for analysis of discourse as a narrative act, with emphasis on actions of characters, time and modality. We also described a similarity metric for calculating inter-annotator agreement. In an initial experiment, two annotators achieved moderate agreement when creating parallel story graphs for 20 fables. In the future, we will grow the Aesop corpus, expand to contemporary genres and perform learning experiments from the data. We will also develop our extension of the model to cover some of the interpretative facets of stories, expanding the user interface and running a new collection experiment that elicits encodings of the inner worlds of characters.

8. Acknowledgments

This material is based on research supported in part by the U.S. National Science Foundation (NSF) under IIS-0935360. 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 NSF.

9. References

- Mieke Bal. 1997. *Narratology: Introduction to the Theory of Narrative*. University of Toronto Press, Toronto, second edition.
- Peter S. Bearman and Katherine Stovel. 2000. Becoming a nazi: A model for narrative networks. *Poetics*, 27:69–90.
- Ofer Biller, Michael Elhadad, and Yael Netzer. 2005. Interactive authoring of logical forms for multilingual generation. In *Proceedings of the 10th European Workshop on Natural Language Generation (ENLG-05)*, pages 24–31, Aberdeen, Scotland.
- Jerome Bruner. 1991. The narrative construction of reality. *Critical Inquiry*, 18:1–21.
- Lynn Carlson, Daniel Marcu, and Mary Ellen Okurowski. 2001. Building a discourse-tagged corpus in the framework of rhetorical structure theory. In *Proceedings of the Second SIGdial Workshop on Discourse and Dialog*, Aalborg, Denmark. ACL.
- David Christian and R. Michael Young. 2003. Comparing cognitive and computational models of narrative structure. liquid narrative technical report tr03-001. Technical report, Liquid Narrative Group, Department of Computer Science, North Carolina State University, Raleigh, NC.
- Jacob Cohen. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1):37–46.
- David K Elson and Kathleen R McKeown. 2007. A platform for symbolically encoding human narratives. In *Proceedings of the AAAI 2007 Fall Symposium on Intelligent Narrative Technologies*, Arlington, VA.
- David K. Elson and Kathleen R. McKeown. 2009. A tool for deep semantic encoding of narrative texts. In *Proceedings of the ACL-IJCNLP 2009 Software Demonstrations*, pages 9–12, Suntec, Singapore.
- Christiane Fellbaum. 1998. *WordNet: An Electronic Lexical Database*. MIT Press, Cambridge, MA.
- Arthur C. Graesser, Kathy Lang, and Richard Roberts. 1991. Question answering in the context of stories. *Journal of Experimental Psychology: General*, 120:254–277.
- Jerry R. Hobbs and Andrew S. Gordon. 2005. Encoding knowledge of commonsense psychology. In *Proceedings of the 7th International Symposium on Logical Formalizations of Commonsense Reasoning*, pages 107–114, Corfu, Greece.
- Paul Kingsbury and Martha Palmer. 2002. From treebank to propbank. In *Proceedings of the Third International Conference on Language Resources and Evaluation (LREC-02)*, Canary Islands, Spain.
- Karin Kipper, Anna Korhonen, Neville Ryant, and Martha Palmer. 2006. A large-scale extension of verbnet with novel verb classes. In *Proceedings of the 12th EU-RALEX International Congress*, Turin, Italy.
- John Lee. 2007. A computational model of text reuse in ancient literary texts. In *In Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics (ACL 2007)*, pages 472–479, Prague.
- Dekang Lin. 1998. An information-theoretic definition of similarity. In *Proceedings of the Fifteenth International Conference on Machine Learning*, pages 296–304, Madison, WI.
- Wook-Hee Min, Eok-Soo Shim, Yeo-Jin Kim, and Yun-Gyung Cheong. 2008. Planning-integrated story graph for interactive narratives. In *Proceeding of the 2nd ACM international workshop on Story representation, mechanism and context*, pages 27–32, Vancouver, British Columbia, Canada.
- Frederick Mostellar and David L. Wallace. 1984. *Applied Bayesian and Classical Inference: The Case of The Federalist Papers*. Springer, New York.
- Erik T. Mueller. 2004. Understanding script-based stories using commonsense reasoning. *Cognitive Systems Research*, 5(4):307–340.
- Martha Palmer, Hoa Trang Dang, and Christiane Fellbaum. 2007. Making fine-grained and coarse-grained sense distinctions, both manually and automatically. *Natural Language Engineering*, 13(2):137–163.
- Martha E. Pollack. 1990. Plans as complex mental attitudes. In Philip R. Cohen, Jerry Morgan, and Martha E. Pollack, editors, *Intentions in Communication*. MIT Press.
- Richard Power, Donia Scott, and Roger Evans. 1998. What you see is what you meant: direct knowledge editing with natural language feedback. In *Proceedings of the 13th European Conference on Artificial Intelligence (ECAI-98)*, Brighton, UK.
- Rashmi Prasad, Nikhil Dinesh, Alan Lee, Eleni Milt-sakaki, Livio Robaldo, Aravind Joshi, and Bonnie Webber. 2008. The penn discourse treebank 2.0. In *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008)*.
- James Pustejovsky, Patrick Hanks, Roser Sauri, Andrew See, David Day, Lisa Ferro, Robert Gaizauskas, Marcia Lazo, Andrea Setzer, and Beth Sundheim. 2003. The timebank corpus. *Corpus Linguistics*, pages 647–656.
- William J Rapaport. 1986. Logical foundations for belief representation. *Cognitive Science*, 10:371–422.
- Mark Riedl and R. Michael Young. 2005. Open-world planning for story generation. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, Edinburgh, Scotland.
- Mark Riedl, Jonathan P. Rowe, and David K Elson. 2008. Toward intelligent support of authoring machinima media content: Story and visualization. In *Proceedings of the 2nd International Conference on Intelligent Technologies for Interactive Entertainment (INTETAIN '08)*, Cancun, Mexico.
- Marie-Laure Ryan. 1991. *Possible Worlds, Artificial Intelligence and Narrative Theory*. Indiana University Press, Bloomington.