Detecting Story Analogies from Annotations of Time, Action and Agency

David K. Elson

Columbia University
New York City

Abstract

We describe the Story Intention Graph (SIG) as a model of narrative meaning that is amenable to both corpus annotation and computational inference. The relations, focusing on time, action and agency, can express a range of thematic scenarios and lend themselves to the automatic detection of story similarity and analogy. An evaluation finds that such detection outperforms a propositional similarity metric in predicting human judgments of story similarity in the Aesop domain.

1. Introduction

Narrative constructs are key to way we perceive, understand and reflect upon information (Bartlett, 1932). We understand stories and events in the context of previous stories we have heard and previous events we have experienced. This can be seen in the many allegories and metaphors that have a narrative basis. A government's austerity measures, for instance, may threaten to "kill the goose that laid the golden eggs," an overly zealous individual may "cry wolf" too many times, or a particularly dangerous turn may take one "out of the frying pan and into the fire."

An algorithm capable of finding structural similarities between stories can greatly assist us in our need to filter, search, and otherwise organize the many stories to which we are exposed on a daily basis, from news articles to fiction and personal communication. Much like a trained language model allows us to recognize n-grams as being more than the sum of their parts, a data bank of encoded stories would let us identify "narrative idioms" that recur and are likely to appear in future stories. To accomplish this, we first need a symbolic model for representing narratives that is sufficiently formal to allow us to algorithmically detect meaningful analogies, yet general enough so that manual tagging of existing stories is feasible (for building the data bank) and automatic tagging is plausible. In a sense, we aim to accomplish automatically the type of structuralist analysis of similarities and trends that Propp (1969) performed on Russian folk-tales and Bremond (1970) on French folktales, using manual annotation as a bootstrap.

We describe the **Story Intention Graph** (or SIG) as a set of discourse relations designed to meet these dual goals. The relations and their adjacent entities can be interpreted as node and arc types in a semantic network, and a particular instance of the SIG model, a "SIG encoding," represents a narrative as a connected graph. To use Formalist terms (Bal, 1997), the SIG captures the underlying sequence of story-world events (the *fabula*) as well as their selection and ordering in the surface rendering of the story (the *sjužet*). Unlike most prior models of narrative discourse that have been proposed for discourse annotation, the SIG has an emphasis on agency. It encodes the links between an action and the intention of its agent (Poynor and Morris, 2003), between a goal-driven action and its outcome (Magliano and Radvansky, 2001), between a goal and its subgoal or superordinate goal (Richards and Singer, 2001), between an event and an affectually impacted agent (Stein et al., 2000), and more. We have used a custom annotation tool to collect a corpus of 70 SIG encodings, collectively called **DramaBank**, from trained annotators.

This paper summarizes the SIG model (Section 3) and its utility in modeling a range of narrative tropes, then presents several approaches to detecting similarities between encodings (Section 4). We compare and evaluate these algorithms against ratings of the similarities among fables, then conclude in Section 5.

2. Related Work

The notion of diagramming narrative as a semantic network is sometimes seen in cognitive psychology (Graesser et al., 1991; Trabasso and van den Broek, 1985). Artificial intelligence originally saw narrative as emerging from scripts, plans, agent interactions or models of common sense (Cullingford, 1981; Wilensky, 1983). Story grammars were also in vogue for a brief period (Prince, 1973; Rumelhart, 1975; Mandler and Johnson, 1977). More recent work in semantic story understanding tends to employ first-order logic (Mueller, 2004; Mueller, 2006; Zarri, 2010) and other formal representations of plans and strategies (Hobbs and Gordon, 2005). Story generation presents its own unique challenges (Gervás et al., 2006) but can also use a planning framework (Riedl and Young, 2005).

Some recent studies have striven to find statistical patterns in corpora of narrative discourse (Chambers and Jurafsky, 2008; Gordon and Swanson, 2009) or build classifiers that adopt Lehnert's (1981) notion of recombinable *plot units* as a discourse model (Appling and Riedl, 2009; Goyal et al., 2010; Nackoul, 2010). Our SIG model bears some similarities to that of the plot unit, but has a greater expressive range by adopting a "theory of mind" approach to literature (Palmer, 2007). This emphasis on the internal states of discrete, intentional agents is also featured in Wiebe's (2005) model of private state frames, as well as Grosz and Sidner's (1986) representation of speaker intention and a recent computational treatment by Chen and Fahlman (2008). Other recent work has adopted the theory-of-mind approach to reading a text, with its emphasis on epistemic differences between agents, in order to model real-life narratives (Löwe et al., 2009; Nissan, 2008). Our initiation of a "DramaBank" collection runs parallel to the "StoryBank" approach of Finlayson (2008); the latter involves the broad annotation of many aspects of sentential-level discourse coherence, where we focus on time, modality and agency (as either stated or implied by the receiver).

The problem of detecting and generating analogies has a long history as well (see (French, 2002) for a review), but not traditionally in the narrative sense. When attempted, such as by Winston (1980) and Finlayson (2009), a method of narrative analogy detection is sensitive to the choice of representation used (Löwe, 2010). It is safe to say that a crucial aspect of any approach to analogy detection is the design of the representation. As the current inquiry is no exception, we emphasize the SIG as a means for describing meaningful temporal and agentive relationships among stories.

3. Story Intention Graphs

The SIG is a constructionist model, in that it brings out coherence at both local and global levels: what happens, when, why, and to whom. In each encoding, a discourse is connected to a representation of its meaning in a single, integrated graph.

In the annotation process, the discourse is divided into fragments, typically of clause or sentence length. Each fragment is represented by a **Text (TE)** node. Text nodes are chained together by **followed by (f)** arcs which reproduce the ordering of the fragments in the original discourse. Events, actions and statives that occur in the story-world's underlying *fabula* timeline, as opposed to fragments of the story's telling, are represented in separate entities called **Proposition (P)** nodes. Text nodes connect to equivalent Proposition nodes with **interpreted as (ia)**, and the order of P nodes in the story *fabula* is also determined by **followed by** arcs.[1] This dichotomy allows us to represent temporal disfluencies such as flashbacks. P nodes are annotated with discrete **agents**, and may also contain fuller encodings of their textual equivalents in any format (such as predicate-argument structures).

The remaining "interpretative" nodes and arcs describe a reader's cumulative situation model (Zwaan and Radvansky, 1998) over the course of comprehending the entire narrative, including both stated and inferred content. In this context, an event, action or stative is represented by an **Interpretative Proposition (I)** node. A **Belief (B)** node acts as a frame containing content that represents a particular agent's conception of the story-world. A **Goal (G)** node is similar to a Belief, except that the nodes and arcs inside a Goal frame are understood to be the state of the story-world as desired by the discrete agent. P nodes connect to interpretative frames and nodes through six arcs:

- **interpreted as (ia)**, **implies (i)** and **actualizes (ac)** are "actualizing," in that they indicate a positive functional relationship;
- **prevents/ceases (pc)** indicates a negative functional relationship;
- **attempt to cause (ac)** and **attempt to prevent (ap)** indicate agent intention to either actualize or prevent/cease.

The first four differ in their directness: *Interpreted as* indicates direct equivalence, *implies* indicates obvious entailment; *actualizes* indicates a positive but indirect causal re-

[1]This is a notational abbreviation for the SIG's interval model of *fabula* time, in which P nodes relate to State nodes in a series of timelines. A concurrent conference paper (Elson, 2012a) gives a fuller description of the schemata.

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."
A Lion watched a fat Bull feeding in a meadow, and his mouth watered when he thought of the royal feast he would make, but he did not dare to attack him, for he was afraid of his sharp horns.
Hunger, however, presently compelled him to do something: and as the use of force did not promise success, he determined to resort to artifice.
Going up to the Bull in friendly fashion, he said to him, "I cannot help saying how much I admire your magnificent figure. What a fine head! What powerful shoulders and thighs! But, my dear friend, what in the world makes you wear those ugly horns? You must find them as awkward as they are unsightly. Believe me, you would do much better without them."
The Bull was foolish enough to be persuaded by this flattery to have his horns cut off; and, having now lost his only means of defense, fell an easy prey to the Lion.

Table 1: "The Fox and the Crow" (top) and "The Wily Lion", from Jones (1912).

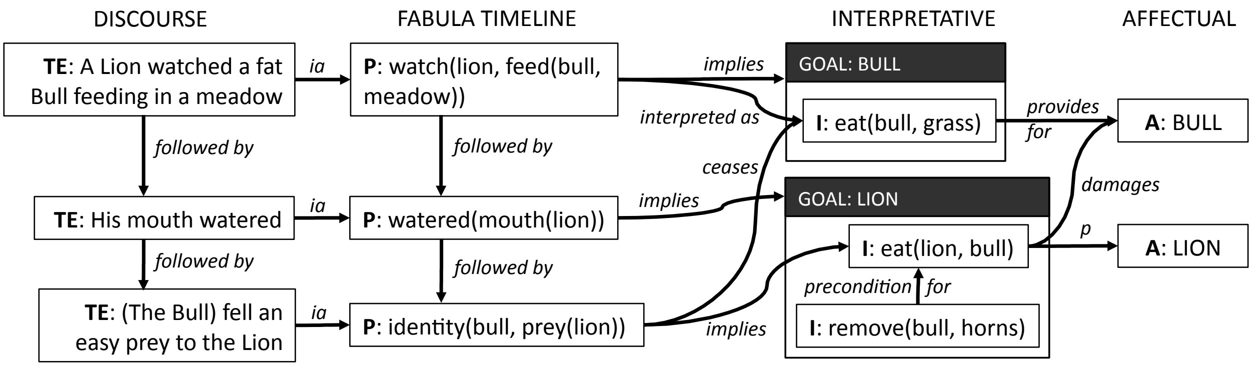


Figure 1: Example SIG encoding for a non-contiguous fragment of "The Wily Lion".

lationship; and *prevents/ceases* indicates a negative, indirect causal relationship.

An example SIG encoding for part of the Aesop fable "The Wily Lion" (Table 1) is shown in Figure 1. Three TE nodes contain text spans and are connected to three P nodes with equivalent propositions. There are two interpreted goals in this encoding: The bull has a goal to eat grass (his will to live is implied), and the lion has a goal to eat the bull. Note that frames themselves refer to mental states: The first sentence implies that the bull has a desire to eat grass, and directly asserts that it is, in fact, eating grass, such that he begins the story with a satisfied goal; later, the "goal content" of eating grass is ceased when the lion eats the bull. This event coreference—the same action is desired, achieved and then lost—forms the basis of our approach to modeling narrative cohesion.

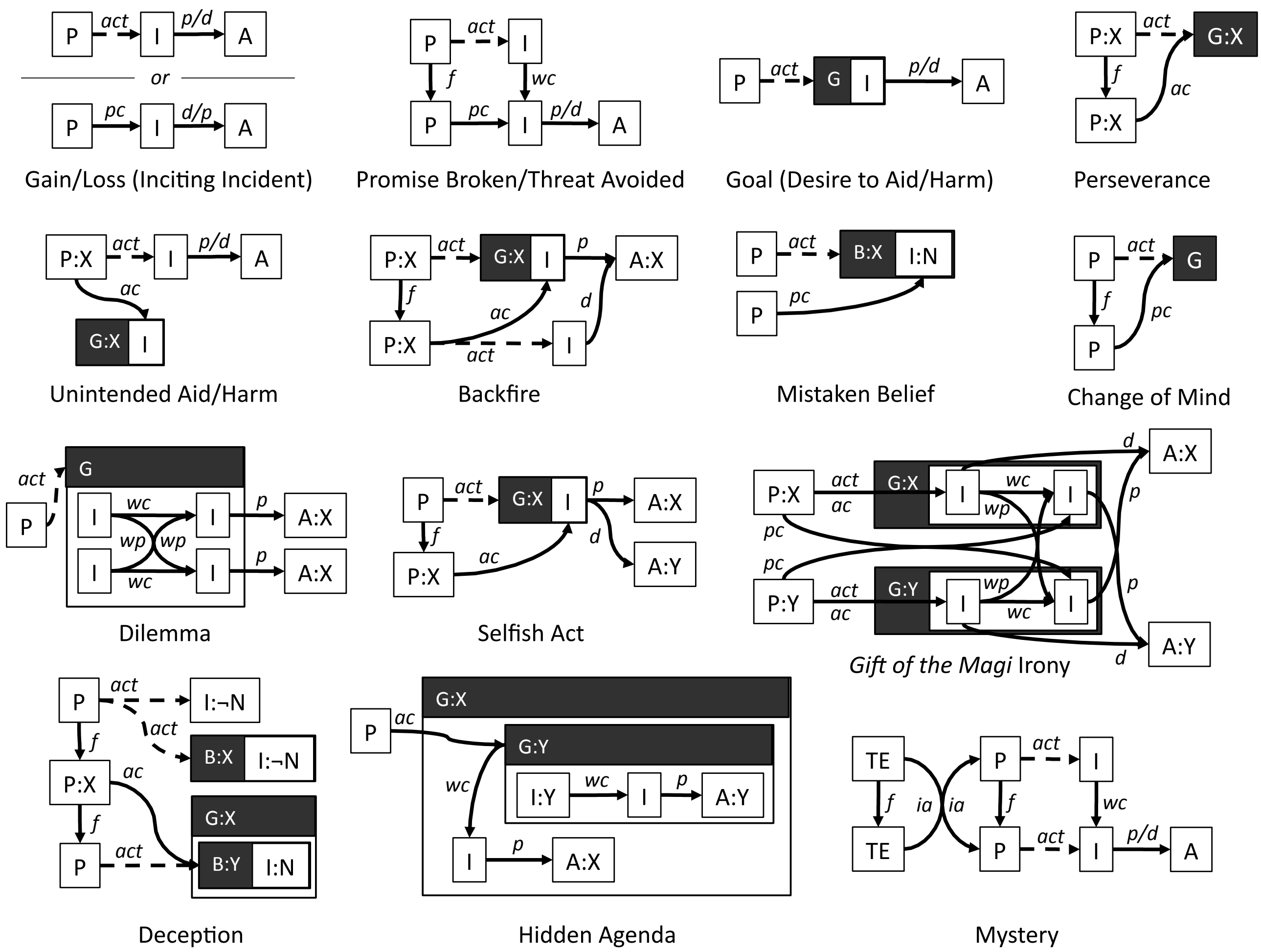
We indicate an agent's plan toward reaching a goal by placing additional nodes of goal content inside the goal frame, and linking them into a causal chain. Each goal becomes a "subgoal" whose actualization is necessary for the superordinate goal that follows (though the relationship is itself a belief on the agent's part). Specifically, a **would cause (wc)** relation links one interpretative frame or proposition to another interpretative frame or proposition, and means that in the belief context of the originating node, an actualization of the first node is both necessary and sufficient for the actualization of the second node. We conversely use **would prevent (wp)** to represent a belief that the actualization of the first node would cause the second node to be prevented/ceased. To signify a belief that actualizing one node is necessary but not sufficient for actualizing or preventing/ceasing another, we use **precondition for (pf)** and **precondition against (pa)**, respectively. In both of the fables we are considering, an agent devises a plan in which one step is for a second agent to devise a plan of its own. For instance, the lion schemes for the bull to act in pursuit of the bull's ego; we would place the bull's goal frame inside the lion's.

Finally, the affectual impact of a P node or actualized I node can be indicated through the combination of an **Affect (A)** node (with respect to a particular agent) and a **provides for (p)** or a **damages (d)** arc (which indicate positive and negative impact, respectively). In Figure 1, the bull's eating of grass is intrinsically good for the bull, while the lion's eating of the bull is good for the lion but bad for the bull. In a properly formed SIG encoding, every goal is annotated with its affectual impact either through a direct arc or indirectly through a plan. The bull's removing its horns, for instance, is indirectly good for the lion because it satisfies part of the lion's plan.

This set of nodes and arcs forms a basic vocabulary and syntax from which complex narrative structures can be constructed. This can be seen through the enumeration of "SIG patterns," compound relations that serve as fragments of abstract narratives. We intuitively define a set of *a priori* SIG patterns to represent a range of narrative scenarios, in a manner similar to Lehnert's enumeration of plot units but with a greater emphasis on temporal and agentive (theory-of-mind) relationships. Notably, these patterns are defined only in terms of node and arc permutations, without any notion of particular propositional content within P and I nodes (except to identify the agent, such as P:X for agent X). We have identified 80 patterns (Elson, 2012b), fourteen of which are shown in Figure 2, in several categories: affectual status transitions (e.g., gain, loss, mixed blessing), single-agent goals (problem, obstacle), outcomes (backfire, lost opportunity, recovery, peripeteia), beliefs (surprise, anagroisis, false dawn), dilemmas, two-agent interactions (selfless act, conflict, coercion, betrayal), persuasion and deception (mutual deception), time (flashback, suspense), mystery, and contradictory points of view. The dotted arcs labeled *act* represent any of the arcs that actualize (*ia*, *i* or *ac*). This is not an exhaustive list of possible SIG patterns, but rather a demonstration of the range of tropes that such patterns can express.

The DramaBank collection project, underway and publicly available,[2] elicits SIG encodings for stories in various genres from trained annotators. Each machine-readable

[2] http://www.cs.columbia.edu/~delson



Pattern	Example	Pattern	Example
Gain	John made a sale.	Change of Mind	Oscar briefly took up the violin.
Promise Broken	The train arrived, but skipped the station.	Dilemma	Betty wanted to be both a full-time chef and a full-time mom.
Goal	Mary dreamed of being published.	Selfish Act	Zach refused to give the old lady his seat on the bus.
Perseverance	Phil courted Megan for years.	*Gift of the Magi* irony	Della sold her hair to buy a chain for Jim's watch, but in the meantime, Jim sold his watch to buy Della a set of beautiful combs.
Unintended Harm	Lou's party, while fun, helped to spread a nasty flu.	Deception	Paul gave a check to the jeweler that he knew would bounce.
Backfire	Francis argued for a better grade, but annoyed his teacher into a deduction.	Hidden Agenda	The fox challenged the crow to demonstrate her singing ability, so that she would drop a piece cheese that the fox desired.
Mistaken Belief	It was clear out, but Yaël thought it was raining.	Mystery	Hillary jumped out of the burning building. She was performing a stunt for an action movie.

Figure 2: Fourteen examples of SIG patterns, compounded relations that represent common narrative scenarios.

record includes a reproduction of the source text as well as the nodes and relations of the annotator's encoding (serialized as first-order predicates). The collection includes 60 encodings covering 25 of Aesop's fables (Jones, 1912), as well as 10 encodings covering 8 samples of longer and more varied narrative discourse: a news article (*Wall Street Journal*), literary short fiction ("An Alcoholic Case" by F. Scott Fitzgerald, "The Gift of the Magi" by O. Henry, and "The Lady with the Dog" by Anton Chekhov), contemporary nonfiction (an excerpt from *Sled Driver*, by Brian Shul), and epic poetry (*Beowulf* and *The Battle of Maldon*). For the 60 Aesop encodings, annotators supplied precise propositional content in P and I nodes according to a controlled vocabulary of nouns, verb frames, and modifiers (Elson and McKeown, 2009). For the longer and more complex texts, annotators constructed SIG encodings that only indicated an agent for each P and I node in order to accelerate the process. Further details on the collection process appear in Elson and McKeown (2010; 2012a).

4. Analogy Detection

Using this model as our representation, we define an analogy among narratives as a SIG fragment that is *covered* by part or all of the SIG encodings of two or more constituent stories. An encoding that covers a second encoding has a subgraph that isomorphic to the graph structure of the second encoding. For instance, if two encodings both feature a Proposition which is *interpreted as* a Goal frame containing an Interpretative Proposition, which *provides for* an Affect node, both encodings cover the "Desire to Aid" pattern in Figure 2, and thus the stories are analogous in that they both involve an abstract character with a desire to aid itself or another agent.

When two I or P nodes are found to be counterparts (analogous) within the isomorphism, we can also compare the propositions themselves using hypernym trees that correspond to each predicate and argument in the controlled vocabulary. As we describe in Elson and McKeown (2010), the analogous proposition would feature the least general predicates and arguments that are hypernyms to both of the constituent propositions. "A Lion watched a fat Bull," for example, would match "A Fox observed a Crow" with the generic "An animal perceives a second animal," and a scoring heuristic would judge this to be a fairly close match (a strong analogy). Our prior work used this technique alone, without any graph isomorphisms except for temporal sequencing, to find story analogies; here, we use this **propositional similarity** algorithm as a baseline approach.

This section ignores propositional similarities in exploring two methods for detecting story analogies based on isomorphisms alone: **static pattern matching**, a top-down approach, and **dynamic analogy discovery**, which is bottom-up.

4.1. Static pattern matching

In the previous section, we described a subset of 80 SIG patterns that we identified *a priori* (without examining the DramaBank encodings). These express narrative tropes in terms of SIG relations. We can apply these as features in a metric of pairwise analogical distance, in that more analogous stories will cover more SIG patterns in common.

The first step is to define and apply a set of logical closure rules. These determine the extent of the transitive arcs that can be derived from annotated arcs. For instance, an *attempt to cause* an event which would have a *positive* affectual impact on some agent should be equivalent to an *attempt to prevent* an actualized event which has a *negative* impact, as both are essentially attempts to effect a net positive change for the agent in question. The closures we have identified allow analogies to be detected despite minor variations in graph structure. Once the transitive arcs are in place, we use a theorem prover (Prolog) to determine whether either of the stories in question covers each pattern at least once, compile two vectors from these 80 features and calculate their cosine similarity.

As a baseline check for the validity of this approach, we leverage the fact that DramaBank contains 60 encodings of 26 unique fables, including 40 homogeneous pairs of encodings (same source story, different annotators) and 1,015 heterogeneous pairs (different stories). We would naturally expect the similarity scores for homogeneous pairs to be significantly higher than those for heterogeneous pairs—while we expect differences between parallel encodings of the same stories, given the subjective nature of story understanding and the flexibility of the SIG model, these differences should not exceed those between opposing stories. We do, in fact, find this to be the case: By the two-tailed Student's t-test, homogeneous pairs are more similar to $p<.001$. One downside to static pattern matching, though, is that the vocabulary of possible analogical overlaps is limited to the set of patterns that we have provided. This method cannot describe specific analogical connections between encodings.

4.2. Dynamic analogy discovery

A third approach to analogy detection finds the largest isomorphic subgraph between two encodings in a manner that respects the semantic constraints of the model; in effect, this finds the most complete and detailed continuous chunk of overlap between two stories (limited, of course, to those overlaps which can be expressed by the model's relations).

We model our algorithm after the ACME model (Holyoak and Thagard, 1989) for finding analogies in connectionist networks. After applying the same transitive closure rules to each encoding, we first seed a set of small "globs" that represent potential isomorphisms between two encodings, then grow each glob by following outgoing corresponding arcs to identify and add new analogous node pairs. That is, if each node in a certain node pair connects to an unseen node via the same relation, the adjacent nodes are paired together. Every glob contains a *binding* which lists not only the discovered node pairs, but pairs of analogous agents as well—as the glob grows, the agent bindings must remain consistent for the analogy to be valid. If agent X in one story is bound to agent Y in a second story, the glob cannot expand to include a node pair in which X would bind to an agent other than Y.

The seeding process begins by considering all possible analogical node pairs among the interpretative nodes (goals, plans and beliefs). If there are conflicting node pairs to which a singe glob can expand, the glob forks into two to track both alternatives. To avoid intractable growth, aggressive memoization is used so that we avoid considering the same glob twice.

Once a glob has expanded to the point where no additional node pairs can be added, we determine which pairs of P nodes in the *fabula* timeline would be consistent with its binding, then add as many P-node pairs as possible by using the Needleman-Wunsch (1970) alignment algorithm. This efficiently finds the longest path of P-node pairs that is internally consistent and compatible with the glob binding.

The result at this point is a set globs that relate to different parts of the agentive content (multiple disjoint isomorphisms). We combine as many as possible into a final analogy by examining each glob in descending order of size, and merging it into the largest glob with which it has a compatible binding. When this is complete, our final result is a set of mutually incompatible analogical bindings that align not only timeline propositions, but agentive content found to be isomorphic between the two encodings. We give each glob a score by counting the relations, nodes and agents found to be analogous in its binding. The top-scoring (largest) glob becomes the top-line result, a particular analogical overlap between two encodings.

Results

We have found this algorithm to return substantive analogies, as measured by the sizes of the isomorphic subgraphs that are found: 8.8 bound node pairs, 1.5 agentive bindings and 14.1 analogous relations on average (including inferred, transitive relations) across 1,015 heterogeneous encoding pairs in DramaBank. We also find again that homogeneous pairs yield significantly larger analogies than heterogeneous pairs ($p<.001$), by more than 50%.

The largest analogies found in the corpus, by the number of bound node pairs, were between two particular encodings of "The Wily Lion" and "The Fox and the Crow". This is an initial check on our approach, as while we did not develop the algorithm using this or any particular pair of encodings, we did include these fables in the collection due in part to their strong analogical connection. By drawing each bound node pair as a single compound node, we visualize this analogy as a hybrid encoding in Figure 3. In this case, there are 11 aligned timeline propositions, two goal frames (one nested within the other as part of a four-stage plan), and two Affect nodes. The overall result is that "the fox is like the lion" and "the crow is like the bull"—in both stories, one is an inciting agent who devises a plan to have a victim devise and execute its own plan that would benefit the inciter. After some persuasion, the inciter's plan succeeds.

4.3. Evaluation Against Gold Standard

In order to evaluate whether we are finding meaningful analogies with each approach, we conducted an evaluation to determine the extent to which we can approximate human ratings of story similarity.

Using Mechanical Turk, we presented raters with each pair of Aesop fables among the 26 we collected, and asked them to rate the degree of similarity on a three-point Likert scale. Our prompt asked for "similarities about story structure and content, such as similarities in plots (what happens) and characters (desires and personality traits)." We presented each story pair to three annotators. The unanimous agreement on the Likert question was 46.3%, with another 50.4% of cases showing a two-to-one majority. To control for nonsense input (as is always a concern with Mechanical Turk), we identified and discounted those individuals whose rate of participation in unanimous agreement was less than 20%; this affected 3.9% of the total vote count. We took the arithmetic mean of the ratings for each pair as its canonical similarity.

We then trained a linear regression model on 100 predictor variables separated into three sets, one for each of our three similarity metrics. Variables regarding propositional similarity included the number of overlapping propositions between the two encodings and the closeness of the overlaps. Each of the 80 static SIG patterns was included as a variable. For the dynamic analogy metric, we included various features relating to the largest detected analogy: number of node pairs, number of agent bindings, types of relations found, and so on. These distributions were normalized and fit against the similarity ratings using M5 attribute selection, and evaluated using cross-validation. We ran the evaluation for all combinations of variable sets to gauge the relative impact of each.

The results are shown in Table 2. Propositional overlap variables by themselves were weak predictors of story similarity ratings, as compared to the other two sets, with an R-squared value of only 0.06. The variables regarding static SIG patterns and dynamic analogies were highly influential by comparison, with R-squared values exceeding .20; the combination of all variables yielded a model which predicted similarity ratings at R^2=.33. This model makes progress toward the prediction of story similarity, with an F-statistic of $p<.0001$. The root-mean-square error is .19,

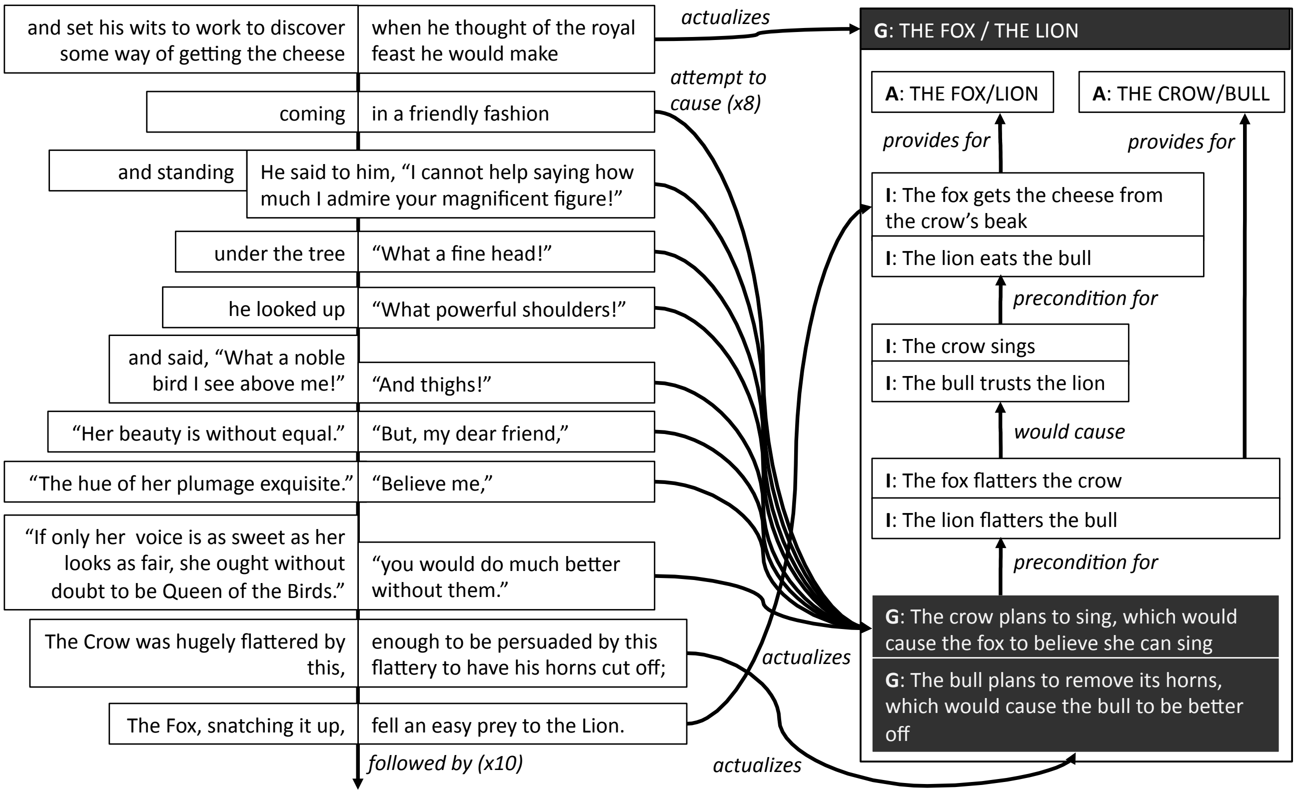


Figure 3: Analogy procedurally discovered between encodings of "The Fox and the Crow" and "Wily Lion".

Predictor Variable Sets	R-Square	RMSE	F-Statistic
Propositional (P)	0.0551	0.1986	p<.0191
Static (S)	0.2729	0.1923	p<.0001
Dynamic (D)	0.2117	0.1948	p<.0001
P+S	0.2724	0.1924	p<.0001
P+D	0.2174	0.1947	p<.0001
S+D	0.3257	0.1893	p<.0001
P+S+D	0.3299	0.1891	p<.0001

Table 2: Cross-validated performance of various linear regression models against story similarity ratings.

compared to .20 for the model with only propositional predictors. In fact, we note that the model including all but propositional predictors performed virtually as well as the all-inclusive model, as measured by both R-squared and RMSE. Propositional modeling, while labor-intensive, did not provide helpful returns on the story similarity task.

The largest caveat of these results is the particularly lopsided distribution of similarity ratings—to most raters, nearly all story pairs had few to no appreciable similarities. Only 99 of these encoding pairs, less than 10%, were rated above 0.5. An increase in the amount of training data, or an expansion of the raters' notion of story similarity, would create a smoother distribution for training our models.

5. Conclusion

We have described a novel set of discourse relations intended to model narrative in a manner suitable for both corpus annotation and algorithmic treatment, for purposes of detecting tropes, similarities and analogies across multiple encodings. The SIG model, featured in a collection of 70 encodings of narratives in various genres, represents not only narrated events, with their temporal and modal relationships, but agents, goals, plans, beliefs, attempts, outcomes and affectual impacts, whether stated or inferred. These relations can be permuted to abstractly describe a range of common narrative tropes. We also described three approaches to detecting analogies, and found that the top-down and bottom-up techniques that leveraged the model's relations outperformed a baseline of propositional similarity against human ratings of story similarity, suggesting that the relations correspond meaningfully to the analogy retrieval task. In future work, they may also lend themselves to a generative model, trained on DramaBank encodings, of story *fabula* and its telling in discourse.

6. 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 author and do not necessarily reflect the views of the NSF.

7. References

D. Scott Appling and Mark O. Riedl. 2009. Representations for learning to summarize plots. In *Proceedings of the AAAI Spring Symposium on Intelligent Narrative Technologies II*, Palo Alto, California.

Mieke Bal. 1997. *Narratology: Introduction to the Theory of Narrative*. University of Toronto Press, Toronto, second edition.

Frederic C. Bartlett. 1932. *Remembering: a study in experimental and social psychology*. Cambridge University Press, Cambridge.

Claude Bremond. 1970. Morphology of the french folktale. *Semiotica*, 2(3):247–276.

Nathanael Chambers and Dan Jurafsky. 2008. Unsupervised learning of narrative event chains. In *Proceedings of the 46th Annual Meeting of the Association of Computational Linguistics (ACL-08)*, pages 789–797, Columbus, Ohio.

Wei Chen and Scott E. Fahlman. 2008. Modeling mental contexts and their interactions. In *Proceedings of the AAAI 2008 Fall Symposium on Biologically Inspired Cognitive Architectures*, Arlington, Virginia.

M. Cullingford. 1981. Sam. In R. Schank and C. Riesbeck, editors, *Inside Computer Understanding: Five Programs Plus Miniatures*, pages 75–135. Erlbaum, Hillsdale, New Jersey.

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.

David K. Elson and Kathleen R. McKeown. 2010. Building a bank of semantically encoded narratives. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC 2010)*, Malta.

David K. Elson. 2012a. Dramabank: Annotating agency in narrative discourse. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC 2012)*, Istanbul, Turkey.

David K. Elson. 2012b. *Modeling Narrative Discourse*. Ph.D. thesis, Columbia University, New York City.

Mark Alan Finlayson. 2008. Collecting semantics in the wild: The story workbench. In *Proceedings of the AAAI 2008 Fall Symposium on Naturally-Inspired Artificial Intelligence*, Arlington, Virginia.

Mark Alan Finlayson. 2009. Deriving narrative morphologies via analogical story merging. In B. Kokinov, K. Holyoak, and D. Gentner, editors, *New Frontiers in Analogy Research*. NBU Press, Sofia.

Robert M. French. 2002. The computational modeling of analogy-making. *Trends in Cognitive Sciences*, 6(5):200–205.

Pablo Gervás, Birte Lönneker-Rodman, Jan Christoph Meister, and Federico Peinado. 2006. Narrative models: Narratology meets artificial intelligence. In *Proceedings of Satellite Workshop: Toward Computational Models of Literary Analysis, 5th International Conference on Language Resources and Evaluation (LREC 06)*, pages 44–51, Genoa, Italy.

Andrew S. Gordon and Reid Swanson. 2009. Identifying personal stories in millions of weblog entries. In *Proceedings of the Third International AAAI Conference on Weblogs and Social Media*, San Jose, California.

Amit Goyal, Ellen Riloff, and Hal Daumé III. 2010. Automatically producing plot unit representations for narrative text. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing (EMNLP 2010)*, Cambridge, Massachusetts.

Arthur C. Graesser, Kathy Lang, and Richard Roberts. 1991. Question answering in the context of stories. *Journal of Experimental Psychology: General*, 120:254–277.

Barbara J. Grosz and Candace L. Sidner. 1986. Attention, intentions, and the structure of discourse. *Linguistics*, 12(3):175–204.

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.

Keith J. Holyoak and Paul Thagard. 1989. Analogical mapping by constraint satisfaction. *Cognitive Science*, 13:295–355.

V. S. Vernon Jones. 1912. *Aesop's Fables: A New Translation*. Avenel Books, New York.

Wendy G Lehnert. 1981. Plot units and narrative summarization. *Cognitive Science: A Multidisciplinary Journal*, 5(4):293–331.

Benedikt Löwe, Eric Pacuit, and Sanchit Saraf. 2009. Identifying the structure of a narrative via an agent-based logic of preferences and beliefs: Formalizations of episodes from csi: Crime scene investigation. In Michael Duvigneau and Daniel Moldt, editors, *Proceedings of the Fifth International Workshop on Modelling of Objects, Components, and Agents*, pages 45–63.

Benedikt Löwe. 2010. Comparing formal frameworks of narrative structures. In *Computational Models of Narrative: Papers from the 2010 AAAI Fall Symposium*, Menlo Park, California.

Joseph P. Magliano and Gabriel A. Radvansky. 2001. Goal coordination in narrative comprehension. *Psychonomic Bulletin & Review*, 8(2):372–376.

Jean M. Mandler and Nancy S. Johnson. 1977. Remembrance of things parsed: Story structure and recall. *Cognitive Psychology*, 9(1):111–151.

Erik T. Mueller. 2004. Understanding script-based stories using commonsense reasoning. *Cognitive Systems Research*, 5(4):307–340.

Erik T. Mueller. 2006. Modelling space and time in narratives about restaurants. *Literary and Linguistic Computing*, 4.

David Douglas Nackoul. 2010. Text to text: Plot unit searches generated from english. Master's thesis, Massachusetts Institute of Technology.

Saul B. Needleman and Christian D. Wunsch. 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology*, 48(3):443–453.

Ephraim Nissan. 2008. Nested beliefs, goals, duties, and agents reasoning about their own or each other's body in the timur model: A formalism for the narrative of tamerlane and the three painters. *Journal of Intelligent and Robotic Systems*, 52:515–582.

Alan Palmer. 2007. Universal minds. *Semiotica*, 165(1–4):202–225.

Daivd V. Poynor and Robin K. Morris. 2003. Inferred goals in narratives: Evidence from self-paced reading, recall and eye movements. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 29(1):3–9.

Gerald Prince. 1973. *A Grammar of Stories: An Introduction*. Mouton, The Hague.

Vladimir Propp. 1969. *Morphology of the Folk Tale*. University of Texas Press, second edition. Trans. Laurence Scott. Originally published 1928.

Eric Richards and Murray Singer. 2001. Representation of complex goal structures in narrative comprehension. *Discourse Processes*, 31:111–135.

Mark Riedl and R. Michael Young. 2005. Story planning as exploratory creativity: Techniques for expanding the narrative search space. In *Proceedings of the 2005 IJCAI Workshop on Computational Creativity*, Edinburgh.

David Rumelhart. 1975. Notes on a schema for stories. In D.G. Bobrow and A. Collins, editors, *Representation and Understanding: Studies in Cognitive Science*, pages 231–236. New York Academic Press, Inc.

Nancy L. Stein, Tom Trabasso, and Maria D. Liwag. 2000. A goal appraisal theory of emotional understanding: Implications for development and learning. In Michael Lewis and Jeannette M. Haviland-Jones, editors, *Handbook of emotions (2nd ed.)*, pages 436–457. Guilford Press, New York.

Tom Trabasso and Paul van den Broek. 1985. Causal thinking and the representation of narrative events. *Journal of Memory and Language*, 24:612–630.

Janyce Wiebe, Theresa Wilson, and Claire Cardie. 2005. Annotating expressions of opinions and emotions in language. *Language Resources and Evalution (formerly Computers and the Humanities)*, 39(2-3):165–210.

Robert Wilensky. 1983. Story grammars versus story points. *Behavioral and Brain Sciences*, 6:529–623.

Patrick H. Winston. 1980. Learning and reasoning by analogy. *Communications of the ACM*, 23:689–703.

Gian Piero Zarri. 2010. Representing and managing narratives in a computer-suitable form. In *Representing and Managing Narratives in a Computer-Suitable Form*, Arlington, Virginia.

Rolf A. Zwaan and Gabriel A. Radvansky. 1998. Situation models in language comprehension and memory. *Psychological Bulletin*, 123(2):162–185.