Chapter 3

Story Intention Graphs

In Chapter 2, we demonstrated methods for identifying a particular type of relationship in narrative discourse (conversational networks). We showed this to be a descriptive aspect of a corpus of Victorian novels. A story, though, cannot be completely described by its social structure alone. There is more that separates stories from non-stories, as not every document with quoted speech is necessarily a story (transcripts of meetings and email exchanges being two counterexamples).

We now turn our attention to a separate discourse model, one that broadens our view of “storiness” beyond quoted speech. We first ask: What kind of model can best capture the essence of a story? What is that essence? We then ask: Can we define a set of intratextual relations specific to our idea of the essence of narrative discourse?

This chapter attempts to answer these questions, first by defining our goals for a narrative representation (Section 3.1), then by reviewing prior models of narrative discourse (Section 3.2), then by describing a new representation (Section 3.3). Our contribution is a set of discourse relations which we collectively call the Story Intention Graphs (SIG). In Chapter 4, we describe the platform and annotation tool we have developed for building and managing a corpus of SIG encodings based on well-known narratives. Finally, Chapter 5 shows that SIGs are an effective formalism for reasoning about stories and their connections to one another.

3.1 Goals For A New Representation

Narrative is the structural scaffold for describing human experience, and so interest in describing narrative structure cuts across a swath of disciplines from the humanities to social sciences and artificial intelligence. People have told stories to one another since before the invention of writing, and some of our earliest recorded thinkers found the proper structure of a story to be worthy of inquiry [Aristotle, 1961]. The notion that a textual narrative invokes any regimented structure at all has swung in and out of fashion repeatedly over the last century among literary theorists, with post-structuralism being a dissenting view. As in any artistic medium, the idea that there are rules that govern “proper” storytelling is an invitation for avant-garde artists to break the rules and further explore the creative space. Still, for centuries, literary theorists have followed Aristotle in describing the form
and content of story, with respect to genre, medium, author, time period, meter, and many other aspects. Modeling narrative is often a means to explore an aspect of the human condition. Some symbolic models, for instance, carry the intention of mimicking our cognitive processes [Graesser et al., 1994]; others use structure as a tool for comparing cultures by their mythologies [Campbell, 1949].

Our purposes are pragmatic by comparison. As we mentioned in the Chapter 1, we aim to find computational methods to look beyond the surface form of a text to compare and contrast stories based on content (as opposed to style). Two stories, even very short ones, may have similar distributions of words at the surface level, and yet convey very different meanings as narrative artifacts. Consider E. M. Forster’s [1900, 87] classic distinction between a non-story and a story:

1. The king died. Then the queen died.

2. The king died. Then the queen died of grief.

Both tell a sequential series of events, both involve the same actions (two deaths), and the word overlap is almost total. While the topic of death is inherently dramatic, (2) is more of a story than (1) because it relates the two events in a coherent manner. It is not simply a matter of attributing causation to the queen’s death; “the queen died of old age” is not a suitable replacement for “the queen died of grief,” which brings to mind images of longing, of a figure bereft to the point of physiological breakdown. Forster also considered this a type of mystery plot: “The queen died, no one knew why, until it was discovered that it was through grief at the death of the king.” Here, as well as in detective stories, the presence of a thematically relevant, unanswered question piques interest in the receiver.

Such facets are examples of what we call thematic content: those qualities that make a story interesting, tellable, memorable, and otherwise evocative of a receiver’s emotional investment. Thematic content is the point that the teller is trying to convey to the reader [Wilenisky, 1982] that separates a story from any chronologically ordered list of events. We search for thematic content when we reflect on a story and think, “What message did it send? Why did I care?” Sometimes, as in poetry, the answer lies closer to the words and the images or patterns they invoke. We are interested in the embedded meaning that persists even when a story is paraphrased, transmuted from one medium to another, or passed from one generation to the next. As we have noted, natural language processing tends to focus on lexical and syntactic features of language; even when artificial intelligence attempts to parse out the deepest meanings of textual stories, the thematic aspect is usually secondary to other concerns [Hidi and Baird, 1986]. With a proper representation, we can identify not only the thematic content of a single story, but the analogical connections between stories: those aspects of thematic content which are common across a group or a genre.

To take a longer example than Forster’s, consider the two fables in Table 3.1: “The Wily Lion” and “The Fox and the Crow”.¹ We intuit that both are stories, in that they

¹The full texts of all of the fables attributed to Aesop that we use in this thesis, as translated by Jones [1912], are reproduced in Appendix D. “The Wily Lion” is fable P469 from the Perry index of these fables [Perry, 2007]; “The Fox and the Crow” is P124.
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.

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.”

\begin{table}[h]
\centering
\begin{tabular}{|l|}
\hline
Table 3.1: “The Wily Lion” (top) and “The Fox and the Crow”. \\
\hline
\end{tabular}
\end{table}

convey meaning beyond the sum of the actions that are described. Both feature discourse relations that set them apart from non-narrative discourse, such as that between an action and the goal that the action is undertaken to fulfill. They are quite analogous: in both cases, a “predator” schemes to obtain something it wants from a “prey” animal who falls for a deception. Both may convey a certain ethical message, advising readers to be conscious of the possibility of ulterior motives when suspect individuals begin to purvey flattery. They take place in a world not totally like our own—ours has no talking lions or foxes—but that does not stop us from seeing how the world is like our own and how the events might somehow relate to our own experiences. We will use “The Wily Lion” repeatedly throughout this chapter as a keystone as we examine prior attempts to develop a descriptive symbolic language for narrative discourse, and then introduce our own.

An approach to finding and describing story analogies like these would provide both intrinsic and extrinsic benefits. Intrinsically, it would shed light on the nature of the stories themselves, as well as on the discourse comprehension processes that allow us to find meaning in them. Extrinsically, a system able to extract narrative themes and tropes from discourse would help us better organize our own narrative and those we experience online or through other media. It could, for instance, let us search news articles, blog posts, and historical literature for all narratives that fit certain parameters—a search that would return entirely different results than one based on keywords. It could find opposing viewpoints on
the same story and help us better understand how, in areas such as politics, two groups can see the same events as belonging to diametrically opposed and difficult-to-reconcile narratives.

Any attempt to build such a system must begin with a complex decision: How should the story be represented symbolically? Which aspects of narrative do we choose to “reify” (formally represent as a narrative primitive) because we expect them to recur from story to story? Any system of symbols implicitly commits to a particular structuralist reading of narrative. If we are to go beyond keyword searching, we must first assert that there is narrative meaning to be found beyond the surface word, and describe a formal representational schemata that captures that meaning. As narrative touches many aspects of logic, language, and culture, this is not a simple task. Let us start by drawing certain boundaries around the space of possible models.

We would like our representation to satisfy three criteria:

- First, we seek a robust model, emphasizing the key elements of a narrative rather than attempting to model the entire semantic world of the story at a high precision. Most of the work in story understanding today occurs in the planning or logic communities; of particular note are the long-term efforts at modeling stories in first-order logic [Mueller, 2004; Mueller, 2006; Zarri, 1997; Zarri, 2010] and other formal representations for plans and strategies [Hobbs and Gordon, 2005; Christian and Young, 2003; Riedl and Young, 2004]. The systems based on these models can compute complex inferences about the story-world and project what can or should happen at some future point in a story. To accomplish this, their models strongly emphasize commonsense knowledge and rule-based inference. Such an approach can come close to addressing the subject matter we are trying to model. In one case, a rule-based system was able to answer questions about the ethics of the stories it read [Reeves, 1991]. It could indicate when people were acting selfishly or selflessly. Mueller's work similarly uses a theorem prover and a database of commonsense axioms to make inferences beyond the first-order assertions that are provided in the story. This is useful for certain why explanations, such as knowing in “The Three Little Pigs” that the house made of straw fell because it was weaker than the house made of bricks.

The drawback of this approach is that the knowledge base tends to be rigid and narrow compared to the information found even in a single novel. No commonsense theorem prover has very wide coverage today, and it is currently infeasible to extract a model of such precision from a discourse automatically. For our purposes, we need to model less about what exactly is happening than why it makes for an interesting story.

This is not to say that we are completely against first-order logic; as we will see, we use this formalism as a way of enriching certain aspects of our representation when they are essential to the thematic content we are trying to capture. For example, our model does not intrinsically understand the differences between ethnic races, but if such a difference is relevant to the discourse relations we are targeting in a particular story (e.g., an important motivation in a story about prejudice), such semantics can be represented within the scope of that story. As Smith and Hancox [2001] noted when describing their own criteria for a narrative representation, different individu-
als can interpret a story at different levels of specificity. Our model will provide a framework for a descriptive first-order representation that allows the precision of the representation to scale up or down as needed. This approach to being tolerant of partial encodings is what we mean by robustness.

• Second, we seek an expressive and computable model of thematic content. As a counterbalance to the first criterion, we wish to find a model that is formal enough to allow us to find analogies, identify patterns, and design summaries of narrative content. A model that uses lexical features to distinguish stories from non-stories [Gordon and Swanson, 2009], for example, is useful for building corpora of stories found online [Swanson and Gordon, 2008], but is only the starting point for distinguishing stories from one another. One recent attempt to find discourse patterns in news stories [Chambers and Jurafsky, 2008a] uses lexical and semantic similarities among words in each story to find “narrative event chains,” in which the same kinds of verbs are applied in the same relative ordering to the same set of protagonists among a collection of articles. While this is a form of analogy, it is specific to news as a genre of narrative—one in which the behavior of the “focalizer” (narrating agent [Genette, 1983; Bal, 1997]) is distinctive. While a journalist typically tries to tell the facts underlying the story transparently, a writer of fiction will strategically withhold, reorder, or obfuscate information for thematic effect. Thus, the key challenge is not just to determine how many relations to extract for representing cross-domain narratives, but to choose the right relations that will strike a balance between wide coverage and semantic precision for this particular task. A story that a reasonable reader would find to have thematically engaging content should have a corresponding encoding under the model that represents those aspects which make the story engaging.

• Finally, we seek a model that is accessible. That is, it should be amenable to manual annotation by human subjects for purposes of building a data bank of narratives. The annotation methodology must be simple and well-documented enough for trained annotators, or even lay readers, to learn. The model should be open-domain, rather than assume a particular input corpus or genre.

Our inspiration here is prior large-scale annotation projects such as the Penn Treebank, which provides syntactic markup of the Wall Street Journal corpus [Marcus et al., 1993], and PropBank, which provides semantic role labeling for individual sentences [Kingsbury and Palmer, 2002]. The Penn Discourse Treebank provides a corpus of documents annotated according to a model of rhetorical relations [Prasad et al., 2008], as does RSTBank [Carlson et al., 2003]. The TimeBank corpus does the same for temporal relations [Pustejovsky et al., 2003b], according to the TimeML model [Mani and Pustejovsky, 2004].

In this spirit, we will present a schemata that we have used to collect a corpus we call DramaBank. Publicly released,\(^2\) it consists of textual narratives annotated with the discourse relations that we find to represent thematic content. It is our hope

\(^2\)http://www.cs.columbia.edu/~delson
that DramaBank will enable further work on narrative discourse parsing, as the Penn Discourse Treebank corpus has enabled recent work on discourse parsers for more expository texts [Lin et al., 2010]. We discuss DramaBank in Chapter 5.

Let us now walk through some of the models of narrative content that have previously been proposed, and consider how well they meet our needs.

## 3.2 A Brief History of Narrative Modeling

We discuss prior narrative models in three categories. First, we consult cognitive psychology and settle on a particular approach to modeling narrative meaning. Second, we review discourse models in general, with an emphasis on story grammars. Finally, we review prior models used in artificial intelligence and natural language processing.

### 3.2.1 Foundations in Cognitive Psychology

The most crucial starting point for teasing apart the meaning of a narrative is the distinction between what the story is and how the story is told. This distinction goes by many names. Todorov [1966] named the plane of story content *histoire* and that of style and point-of-view *discours*. Russian Formalism uses the terms *fabula* and *sjuzhet*, respectively. Genette [1972] refers to the *diegetic level* and *extradiegetic level*. Bal [1997] makes a three-level distinction: the *fabula*, which she defines as “a series of logically and chronologically related events that are caused or experienced by actors” the *story*, in which a narrator (perceiver) selects some elements of the *fabula* to convey and omits others; and the *text*, where words are chosen convey the story in a discourse.” Whichever set of terms is used, the effect is the same. A narrative discourse is a lens that focuses attention on a particular combination of events that transpire in a constructed world. Sometimes the narrator purports that the *fabula* corresponds to our reality, as in a news article by a reliable journalist; in fiction, the world of the story borrows many elements from the actual world but invents others [Eco, 1995].

Much of what happens in a *fabula* remains unsaid, or at least withheld from the reader. Sometimes this is dramatically essential. It would defeat the purpose of a mystery novel for the identity of the killer to be revealed in the first chapter, when the murder takes place. Rather, this crucial fact is omitted by the narrator, who invites the reader to follow the detective’s investigation toward a solution which satisfies all the facts in a causally plausible fashion. Usually, though, these omissions are made for purposes of narrative economy, because they have no bearing on the thematic content of the story. The sentence “John drove to the store,” for example, omits certain elements of its *fabula*: what kind of vehicle John drove, where he was driving from, the time of day, the time of year, the duration of the trip, and so on. When the missing facts are needed to give a story coherence, or interconnectedness as a united discourse, we infer them. For instance, in the Forster citation above, we may infer that the queen died “of grief” in order to connect the two sentences. The thematic content of the story builds off of such inferences, which look beyond the words of the text in search of particular facets of discourse coherence. We strive to understand why the various parts of a story were included in the discourse by the narrating agent, and
how all the parts fit together coherently. When we find such coherence, the story takes on a meaning greater than the sum of its sentences.

But what, in general, are the facets of meaning that we search for? What makes a story more interesting than a set of chronological facts in a fabula? In short, what are the “primitive” symbols of tellable narratives that we can reify in a representation that will allow us to find thematic similarities, differences, patterns and analogies?

Cognitive psychology has been examining these questions in a search for an understanding of the way the mind comprehends discourse. Narrative, as opposed to expository or persuasive discourse, has been a common testbed for understanding the way inferences are generated during reading [Graesser et al., 1991a]. Using methods such as question-answering response time and “think-aloud,” when subjects articulate their inferential processes while reading, they track the way various elements of the story are retained in working memory, discarded, linked together and inferred from background knowledge. As story interpretation is subjective to each receiver, their findings on the process of reading are relevant to the design of our schemata.

Most discourse psychologists subscribe to a three-layer model of discourse representation analogous to Bal’s text-story-fabula system [Graesser et al., 1997]. First proposed by van Dijk, a linguist, and Kintsch, a psychologist [van Dijk and Kintsch, 1983], the model begins with a surface code which (like Bal’s text layer) recalls the words on the page. In cases like poetry where wording is important, this is retained; in news articles and most other forms of narrative, it is quickly discarded in favor of representations of deeper meaning. The second layer, the textbase, contains propositions that convey the essence of what was in the surface code. These are predicate-argument structures, where each argument has a semantic role that relates to the predicate; this is very similar to the first-order representations used in semantic approaches by Muller [2003] and others. This, in turn, is boiled down to a situation model [Zwaan and Radvansky, 1998], which represents the “aboutness” of the text—the states and affairs which are the essence of the story.

As the reader reads, she attempts to integrate each new sentence in the discourse into the situation model using a combination of top-down and bottom-up processes. Psychologists have put forward several theories describing how this is accomplished [Kintsch, 1988; Zwaan et al., 1995]. Most models, though, share a consensus view that a few aspects of a narrative are consistently reified in our cognitive search for meaning: intentional agents and the goals that they would like to see transpire in the story-world.

The evidence for the primacy of these elements reaches into early development. Even before we can read, we see agency in the most abstract of events; separating intention from action is a basic function of narrative perception [Bundgaard, 2007]. In one classic study, Heider and Simmel [1944] showed subjects short animations of geometric shapes moving around a screen. When asked to describe what they had seen, most subjects told stories about animate agents (typically people), the challenges they faced, the love they defended, the assistance they rendered to the needy, and so on. Though the sample size was small, this effect has been experimentally repeated: Subjects easily attribute mental states, involving goals and intentions, to even highly abstract stimuli [Dik and Aarts, 2007]. Other studies show that infants as young as nine months interpret actions around them as causally related to the goals of their agents [Csibra et al., 1999; Gergely et al., 1995]. One study showed that
children as young as six have a better recall of oral stories when characters have well-defined goals [Lynch and van den Broek, 2007]. The children were able to make online inferences as they listened, to connect subgoals to superordinate goals and the actions taken in pursuit of those goals.

These results help explain the large and growing body of evidence in experimental psychology supporting the claim that readers engage in a search for the motivations behind characters’ actions, and are better at retaining and retrieving information about the actions of the story when they are backed by clear goals. In other words, readers are actively searching for goals in order to better comprehend a text [Lichtenstein and Brewer, 1980; Hassin et al., 2005; Aarts et al., 2008]. The nature of the types of inferences and connections that are made, and their timing during the reading process, is the matter of longstanding debate. McKoon and Ratcliff [1992], for instance, advocate a “minimalist” hypothesis in which only those inferences which are necessary for local coherence—sentence to sentence consistency—are made online, and global connections (connecting distant elements of the story) are made only when local coherence is violated by new information. Graesser et al. [1994] propose instead a “constructionist” model in which online inferences include causal antecedents (what caused an action), superordinate goals (what motivates an action) and how a character’s emotional state is affected by an action. They make the crucial point that these findings relate to narrative discourse rather than expository, persuasive or descriptive texts, because narrative is a “privileged” kind of discourse that is closer to the way we perceive and relate everyday experiences [Kintsch, 1980; Bruner, 1986; Nelson, 2003]. We understand ourselves and others partly in terms of overarching goals and the actions we take to pursue them.

Other work has built on these ideas to explore the amount of identification a reader has with the protagonist of a narrative (which allows the reader to adopt the protagonist’s goals as his or her own virtual desires) [Albrecht et al., 1995]. As most stories have multiple agents, often with cross purposes, the goals of different agents must be tracked separately [Magliano et al., 2005]. Suh and Trabasso [1993] find that readers keep both subgoals and superordinate goals in working memory when answering questions about a story, but make only online inferences about the most pressing subgoal during reading, suggesting that they maintain a “hierarchy of goals” in which one must be accomplished in order to achieve the next.

As we have alluded, another consensus view among psychologists is that a basic function of discourse comprehension is to find the relationships between actions and goals. Readers tend to want to view each agent’s actions as being intended in pursuit of one goal or another, to the point where we infer goals when none have been made explicit. Poynor and Morris [2003], for instance, find evidence that goals are inferred at the time the information is presented, even if the information only implies the goal (rather than stating it outright). This suggests that “readers activate a representation of the protagonist and his or her goals early in the narrative, and that representation is strategically maintained throughout the narrative (or at least until that goals is met) as a vehicle for explaining actions.” The import of goals and actions matters as well: Egidi and Gerrig [2006] show that the association between goals and actions is stronger when they are matched in terms of urgency and intensity, respectively.
The constructionist model has also taken on the notion of goal outcome, in which events signal that an agent has either succeeded or conclusively failed to achieve a goal. Researchers have found that this information is also an important trigger in the cognitive indexing of goals, subgoals and actions [Stein and Albro, 1996; Richards and Singer, 2001]. For example, Magliano and Radvanksy [2001] show that the success or failure of a goal affects its prominence in working memory as comprehension continues past the point where the outcome is revealed. Stein et al. [2000] describe the relationship between goals and “emotional understanding,” in that children link happy memories to stories of goal success, and unhappy memories to stories of goal failure and of threats to valued goals. We empathize with the agents in stories and comprehend goal outcomes in terms of affectual impact, identifying agents that are positively impacted by goal success and agents that are negatively impacted by goal failure.

Each of these processes is specific to the reader, and different readers can construct different situation models from the same text. There is no one “true” cognitive representation for a given story. Gerrig and Egidi [2010], for example, point out that not all readers agree on which connections are to be made, and even the same reader can vary in terms of how much inference to bring to a text through periods of reflection. A shallow reader may use reflexive processes and automatic “rules of thumb” to determine the morality of an action or the relative importance of a goal, where a deeper mode of reading triggers alternative connections.

Nonetheless, these results give a basis in cognitive psychology for a set of “narrative primitives” which we use as the basis for our own representation of story logic: intentional agents, goals, subordinate and superordinate goals, outcomes of goals, goal-directed actions, and affectual impacts. A representation that reifies these must also include the basic elements in Bal’s definition of fabula: discrete time events in the story-world and causal relationships between events. Taken together, this set of primitives do well against our three criteria for a representation. Just as experimental subjects can indicate which text spans indicate goals and which are intentional actions in a cognitive study, so too can annotators encode a corpus of stories with similar discourse relations in a computational study. A model based on these elements, aligned with our natural reading processes, would be both robust and accessible. Although most of these findings focus on the comprehension of textual discourse in particular, these primitives are invoked in the situation model regardless of the input modality.

In the remaining part of this section, we will consider three descriptive representations for narrative discourse, and consider the advantages and drawbacks of each.³ In particular, we will examine the way that each system arranges all or some these narrative primitives with a set of useful discourse relations. The three models are:

1. The GRTN (causal network) model, championed by Trabasso as a model of

³We exclude from our survey prior work that uses these primitives to describe aspects of narrative logic, but does not provide a set of discourse relations we might apply to the task at hand. For example, van Dijk’s application of the philosophy of action to the theory of narrative [van Dijk, 1976] considers the logical relationships between actions, intentions, outcomes, and the discourse in which they are expressed. While many of these insights intersect with those of these three models and our own model, we are only describing in detail those approaches which feature an accessible procedure for discourse annotation and analysis.
cognitive story understanding,

2. Linguistically-rooted **story grammars**, in particular the grammar proposed by Mandler and Johnson, and

3. **Plot units**, an influential model that originated in Lehnert’s work in artificial intelligence.

We will continue to use “The Wily Lion” fable as a common point of comparison in considering these three models and setting the scene for our own contribution.

**Recursive Transition (Causal) Networks**

The question of how these narrative primitives relate to one another is the subject of long-standing debate. There is a predominant view that a conceptual network (a “connectionist” model) is constructed in the situation model, but different suggestions for graph topologies have emerged (e.g., [Bearman and Stovel, 2000]). Two suggestions to consider in particular are those of Trabasso and Graesser along with their colleagues.

The “recursive transition” network theory of comprehension proposed by Trabasso [Trabasso and van den Broek, 1985; Trabasso and Sperry, 1985; van den Broek, 1988] organizes the text around the principle of causality. The situation model synthesizes propositions found in the textbase with the reader’s world knowledge to arrive at a graph model in which nodes represent story fragments and edges indicate causation. The model assumes that a traditional story features a protagonist who encounters a problem and goes about a strategy for overcoming it.

The story fragments are separated into six functional categories: **Settings** (S), which establish the protagonist in time and space; **Initiating Events** (E) which result in Internal Reactions (R), which in turn cause the protagonist to have Goals (G); **Actions** (A) which are motivated by goals; and the **Outcomes** (O) of goals. A later articulation of the model [van den Broek, 1988; Trabasso and Wiley, 2005] further distinguishes between successful outcomes (SO) and failed outcomes (FO).

The reader assigns a causal link between any two nodes if there is “necessity in the circumstances,” in that the reader’s world knowledge tells her that one statement is a necessary consequence of the other. A is necessary for B if it is the case that had A not occurred, B would not have occurred. (Since B must temporally follow A to fit this definition, causal networks are also timelines.) Causal links take on slightly different meanings depending on the classes of the nodes they connect (only a “motivation” arc connects a Goal with an Action). Not all nodes can connect to all other nodes—the General Recursive Transition Network (shown in Figure 3.1) indicates which adjacencies are legal.
CHAPTER 3. STORY INTENTION GRAPHS

<table>
<thead>
<tr>
<th>Span</th>
<th>Category</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>S1.1</td>
<td>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:</td>
</tr>
<tr>
<td>2</td>
<td>E1.1</td>
<td>and as the use of force did not promise success, he determined to resort to artifice.</td>
</tr>
<tr>
<td>3</td>
<td>R1.1</td>
<td>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.”</td>
</tr>
<tr>
<td>4</td>
<td>G1.1</td>
<td>The Bull was foolish enough to be persuaded by this flattery to have his horns cut off;</td>
</tr>
<tr>
<td>5</td>
<td>G2.1</td>
<td>and, having now lost his only means of defense,</td>
</tr>
<tr>
<td>6</td>
<td>A2.1</td>
<td>fell an easy prey to the Lion</td>
</tr>
</tbody>
</table>

![Figure 3.2: Outline of “The Wily Lion” in a causal-network representation.](image)

The GRTN model also supports goal hierarchies through goal embedding, which separates the network into a tree of connected causal chains. If one attempt at a goal leads to a failed outcome, for example, the outcome can motivate a subgoal which is understood to be a strategy for achieving the main goal. When these networks are visualized graphically, every node is given a subscript \((i,j)\) where \(i\) is the level of embedding and \(j\) is the node’s position in the order of nodes of its type at its level of embedding. For instance, the second action in the causal chain of a subgoal would be A(2,2).

Figure 3.2 shows a GRTN representation of “The Wily Lion”. First, the table separates the story into spans and assigns a node of a certain category to each span. Below the table is the graphical layout of the network, including causal connections and goal embedding. (The main goal is on the top row; the lion’s subgoal to trick the bull is on the bottom row.) The model brings out many salient aspects of the story using an attractively small number of primitives—six node types and one arc type. It shows that each statement is connected to the main arc of the story, and suggests that some aspects are especially salient to the its structure: those in the critical path from one end of the story to the other, so that they are unavoidable in any trace of the network from S(1,1) to O(1,1). The strong cohesive structure
of the GRTN has been shown to predict the accessibility of causally influential statements in memory [Trabasso and van den Broek, 1985; Trabasso and Stein, 1997]. The model is also robust and formal enough to be the basis for a machine implementation. A simulated version of the process of memory retrieval, using GRTNs, can reproduce the results humans give to recall and “think-aloud” tests on simple stories [Langston and Trabasso, 1999; Trabasso and Wiley, 2005].

There are, though, drawbacks to this type of causal network for the purposes of finding similarities and analogies between stories. For one, the topology describes the point of view of a single protagonist. Many stories find thematic meaning in the competing objectives of multiple agents. For the fable example, the model captures that both “The Wily Lion” and “The Fox and the Crow” involve a goal that begets a subgoal, but the goals of the “prey,” manipulated as they are by their predators, are not included. More generally, causal networks offer a good model of structural importance within a single story but a vague model of the content similarities between stories, even when the structures are similar.

Other psychologists have extended the expressive range of the conceptual graph to address these shortcomings. Graesser and his colleagues [Graesser and Clark, 1985; Graesser et al., 2001] use nine types of arcs instead of one, including Reason, Manner, Consequence and Implies. They also introduce generic knowledge structures (GKSs), which represent aspects of world knowledge, and show how they integrate into the nodes in the graph. Their schemata captures more implied information than Trabasso’s, which is better described as an organization of the textbase around purely causal relationships. We take inspiration from the ability of this network to encode implied information only when it is thematically relevant (as opposed to never, in the case of GRTNs, or as often as possible, as in semantic understanding systems).

Like Trabasso’s graph, Graesser and Clark’s is designed to model human processes for recall, summarization and question-answering, using the constructionist theory of comprehension. We know of no attempt to apply the model to the process of finding analogical connections between different stories, but it has been a fertile testbed for modeling the way inference and activation occur during question-answering. The QUEST process [Graesser et al., 1991b] is able to traverse the conceptual graph to find the best answers to “why” a particular action occurs. (This is often a difficult question to answer, as each action could potentially be explained by any of its causal or motivational antecedents). The graph topology is attractive in its balance of formality and robustness, in that the arcs are carefully selected to cover many stories (those that are totally connected by causal relationships), without being general enough to lump disparate stories into identical graphs. The number of nodes and arcs scales up or down with the number of goals, the complexity of the agents’ plans, and the events that they set in motion. The model, though, is not as accessible as Trabasso’s. Graphs are constructed from stories with a complex verbal protocol in which the researcher questions the subjects, a method which does not scale easily. Still, the QUEST model has found a recent revival in the domain of collaborative co-authoring, where an assistant is able to procedurally ask “why does this happen?” to help the user write a coherent narrative [Riedl et al., 2008].
3.2.2 Discourse and Literary Theory

The linguistic perspective on discourse intersects with that of cognitive psychology. Both aim to find models that connect spans of text by meaningful, functional relations. Just as a syntactic model shows the relations between words that bind the sentence into a functional whole, discourse relations provide cohesion between phrases and sentences to describe the “point” of the discourse. Various types of relations have been proposed to provide coherence to a discourse. Referential relations, for instance, connect multiple mentions of the same entity as they occur throughout a discourse. Coreference and pronoun resolution, the processes that assign referential links, aid discourse comprehension in many languages by connecting clauses and phrases by the entities that are repeatedly mentioned [Grosz et al., 1995; Grishman et al., 2005].

Other discourse relations that have been proposed deal with the way entire clauses and sentences relate to their neighbors. Of particular interest is the manner in which utterances either introduce, expand upon, or break from a topic [McKeown, 1985; Hobbs, 1985; Polanyi and Scha, 1984; Hirschberg and Nakatani, 1996]. Another set of proposed relations describes the structure of implied intentions (what the speaker attempts to accomplish with each utterance) [Grosz and Sidner, 1986; Rambow, 1993]. In this subsection, we focus on those sets of discourse relations which have been proposed to deal with narrative or its close relatives. First, we describe story grammars in the context of other hierarchical models of structure. Then, we summarize other linguistic theories of narrative discourse, and transition to structuralism and other relevant ideas from literary theory.

Hierarchical Models

Expository discourse is the focus of hierarchical models such as Rhetorical Structure Theory (RST) [Mann and Thompson, 1988] and the Penn Discourse Treebank [Prasad et al., 2008]. RST defines a set of relations between text spans. Some of the relations overlap with the narrative primitives we established from the cognitive literature (Volitional Cause, Purpose, Motivation, Sequence). Other relations are geared more toward argumentative discourse (Concession, Otherwise). Each span can subsume sub-spans, resulting in a tree-structured description; Mann and Thompson claim it is “typical, but not universal, for texts to be hierarchically structured and functionally organized.”

Both models have appealing advantages. They strike the balance between generality and formality necessary to lend themselves to large-scale corpus annotation [Carlson et al., 2003; Prasad et al., 2008], corpus-based studies of text organization [Louis and Nenkova, 2010] and automatic relation extraction from surface text [Marcu, 1997; Lin et al., 2010]. However, there are also significant drawbacks. Neither model aligns well with our set of ideal narrative primitives. The relations are not oriented around agents and are not well suited for hypothetical actions (plans and goals). Expository texts, in contrast, are topic-focused, with a model of cohesion that revolves more around the logical chaining of claims and arguments [McCarty et al., 2006; Berman and Nir-Sagiv, 2007].

That is not to say that RST-like models have never been proposed for the narrative mode. For several years in the 1970s and early 1980s, a flurry of work attempted to find a context-free grammar that described the structure of a story [Prince, 1973; Rumelhart, 1975;
Mandler and Johnson's [1977] grammar, reproduced in Table 3.2, is typical. Although Mandler and Johnson presented their work in the context of cognitive psychology, showing the connection between the grammar representation and the results of experiments that measured reader recall, it has influenced linguistic models of discourse such as RST. The grammar's rewrite rules describe the *fabula* of story, including the temporal connections (**AND** for simultaneous events, **THEN** for sequential), causal connections (**CAUSE**), the internal goals of characters, and goal-directed actions. The terminals are **EVENTS** and **STATES** (the latter being a condition of the world), where both can either be internal (as in thoughts, plans, and perceptions) or external (actions and happenings). A simple, goal-directed story like “The Wily Lion” can be parsed by hand to conform to the grammar (see Figure 3.3)—
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.
fact, multiple parses are possible depending on one’s interpretation of causality and other factors that are not made explicit in the text. Note that the terminals in the diagram are sentences from the text, where the grammar calls for more textbase-level terminals (propositions).

Grammars are good for capturing both the local and global coherence of properly structured plots. The quantization of clauses by time states (using the AND and THEN rewrite rules), causal relationships, and associations between goals and goal-driven actions are laid out more precisely than in the non-hierarchical models we saw earlier. Analogical similarities can be found across the corpus by looking for rules and compositions of rules that recur across the corpus. Moreover, it makes intuitive sense that there would be a story-level equivalent for the context-free grammars suggested for syntax.

There are, though, both theoretical and practical problems with using grammars for stories. A context-free grammar that separates stories from non-stories must, by definition, accept every story and reject every non-story. The designer of the grammar must commit to a precise description of what a story is and how it must manifest as surface text. Mandler and Johnson leave some wiggle room—they mention but do not formally describe “transformation rules” that can affect the transition from fabula to discourse, with certain atoms rearranged or deleted—but even at a fabula level, the method is tilted far toward formality at the expense of robustness and coverage. When challenged about the narrow coverage of story grammars [Black and Wilensky, 1979], both Mandler and Johnson [1980] and fellow grammarian Rumelhart [1980] replied that the domain expressed by their grammars was never meant to cover all stories—just those in the oral tradition, or those with a “recursive” structure, respectively. The question remains, though: Can any context-free grammar for stories achieve a wide coverage?

Ryan [1991, 201] argues (before giving her own grammar parse for “The Fox and the Crow”) that the idea of a story grammar is fundamentally suspect because there is a seemingly unbounded number of possible story actions that can serve as terminals. The number of lexemes that can serve as terminals in a syntactic model (i.e., the number of words in the vocabulary) is small compared to the number of actions which we might enumerate as possible in a story-world. More importantly, a fundamental aspect to grammars is that elements in one branch of the parse tree can not “cross over” to relate to neighbors in other branches. Individual terminals can also only belong to a single rewrite rule. Unfortunately, both situations regularly happen in even simple stories, when one interprets them to include the inner worlds of character agency. Characters can revisit and alter plans once they have failed. Two characters may have plans at cross purposes who alter each other’s executions in turns. An undesirable state (such as poverty) can be activated, deactivated and activated again, though we would not suggest having “in and out of poverty” be a high-level rule with its own structural properties. Well-told stories are contrapuntal documents, as seen in popular storytelling advice that writers have each scene fulfill at least two functions at once (a text and a thematic subtext) [McKee, 1997]. Some of this missing information is key to the analogical connections between stories. Like Trabasso’s causal network, which did not itself impose a tree structure on most of its topology, story grammars struggle to see a text from more than one perspective at once. The common experiences of the two victims, the bull and the crow, are enablers for their predators’ plans; this is unfortunate, since the
point of these fables (in the sense of Wilensky’s contemporary story points model [Wilensky, 1983]) is to have the reader identify with the prey. As cautionary tales, they both warn against taking flattery to heart and taking risks in the name of self-aggrandizement.

Similar concerns about the expressive power of trees were later brought up about RST, which, though not strictly a grammar, also provides that only one type of relation can bind two utterances together. As Moore and Pollack point out [1992], in many cases two utterances can be related in multiple ways, even in expository texts. Wolf and Gibson [2005] similarly posit that mandating a tree structure can forego important discourse relations that cross over. Lee et al. [2008], in discussing the shared arguments which are allowed in the Penn Discourse Treebank, ask whether trees are used in discourse simply based on the historical precedent of trees being used widely in syntax. For the case of narrative, we believe that trees are not the right discourse model for finding analogical connections between stories.

Though story grammars in particular fell out of favor, new attempts at this type of model still come up from time to time [Lang, 1999].

Other Linguistic and Narratological Models

Grammars were neither the first nor the last linguistic proposal for modeling narrative discourse (though they were the most controversial). Of interest are other efforts which take a more corpus-based approach to finding a set of useful discourse relations for characterizing group-wide norms by their thematic content.

When linguists Labov and Waletzky [1967] studied a corpus of urban oral storytelling, they found six recurring units: an abstract, which sets the scene; an orientation, which sets the initial situation or activity; a complicating action (what happened?), an evaluation (so what?), resolution, and a coda. Bell [1999] removes some of Labov and Waletzky’s units and adds others to model news articles as narratives, including a premise, a main event, the background that precedes the main event, and so on. Polanyi [1989] considers oral storytelling as a complex social exercise in which the storyteller and her peers take turns and exhibit mutual concern over their images; these pragmatic factors directly influence a story’s content, such as when a controversial position is clarified or qualified.

The group norms do not necessarily have to be about the relations that join together clauses. The content of the textbase can also be the subject of a data-driven study. We can ask not only, “how are stories structurally similar?” but also, “what are stories generally about?” A typical approach to describing this code is to examine an entire corpus and align it by analogically similar content. These normative areas of overlap can be actions that recur in each story, character stereotypes and other tropes. Then, each story is individually examined in terms of how it follows or deviates from the group norms. One recent annotation scheme for analyzing sets of narratives allows semantic units to “emerge” from clusters of text which convey the same basic content, regardless of the degree of lexical overlap [Passonneau et al., 2007]. Another study uses this approach in order to build a system that classifies the completeness of children’s retellings of simple stories [Halpin et al., 2004].

Certainly the most famous example of this technique is Vladimir Propp’s structuralist study of several hundred Russian folk-tales [Propp, 1969]. Propp identified 31 functions
that recur throughout each story, where a function is “understood as an act of a character, defined from the point of view of its significance for the course of the action.” He used his own measure of what we might call “semantic distance” to determine when two events belong in the same function. Each function is given an abstract definition that encompasses all of the scenes it counts as members. Some functions are more abstract than others. The function *The hero is pursued*, for example, is drawn from scenes that span a range of scenarios from the pursuer flying after the hero to the pursuer gnawing at a tree in which the hero takes refuge. Other functions include *The villain causes harm or injury to a member of a family* and *The hero acquires the use of a magical agent*.

Unfortunately, Propp did not give clear guidelines as to *how* he identified each function. That is, he did not tell us exactly how analogous two scenes must be to form a function, so that we can repeat his technique algorithmically. It is also worth noting that not every story supplied an example of every function; some stories contributed to fewer than half. However, Propp’s functions do hold the striking property of sequential consistency, in that they tend to appear in the same order in the corpus (albeit with many exceptions). This allowed Propp to draw up a simple grammar of the Russian folk-tale:

\[ S \rightarrow ABC\uparrow DEFGHJKLMNPQR - RS^0L \]

The letters and up/down arrows represent functions. The fraction indicates a branch point, so that each story follows either the top or bottom path. The rigid ordering of Propp’s functions lends itself to computational treatment, and indeed Propp has widely influenced the field of narrative generation; several studies have even adapted his scheme outright to make story-writing programs [Díaz-Agudo et al., 2004; Gervás et al., 2005]. For the purposes of our task, we find Propp’s method of identifying group norms and their discourse relations appealing, but Propp’s functions are not, in and of themselves, a representation we can apply to Aesop’s fables or other genres outside of Russian folk-tales. We are more interested in replicating the process of identifying group norms that are abstract enough to have wide coverage, yet precise enough to reveal useful insights. Indeed, this is a summary of our larger objective: Our present search for discourse relations that capture thematic content is a search for a representation that enables a systematic search for corpus-wide group norms that describe interesting narratological tropes. Indeed, Propp’s functions are connected only by temporal sequence. We strive to find group norms centered around sets of relations that tie similar events together, including sequence as well as such factors as motivation and enablement that Propp did not model.

Propp was also an inspiration to the mid-century proponents of a structuralist literary theory that sought to find universal system of symbols and relations that organize sentences into coherent narratives. Levi-Strauss [1968] went beyond the story or the function as a unit of analysis, and identified “mythemes” as atomic, irreducible components found across many of the myths of a culture. Like notes forming chords in a musical stave, the mythemes can be joined together (“bundled”) when they co-occur in a story, particularly in binary opposition to one another; these bundles are depicted as a dimension of discourse perpendicular to the temporal flow of time. Todorov [1969] found what he called a “narrative syntax”
underlying several stories in the *Decameron*, using the grammatical structure of language as an analogy. The clauses are connected by temporal ordering, entailment and spatial relationships, similar to the causal networks proposed in psychology. To Barthes [1975] as well, discourse is a large “sentence,” with a grammar that links units together into a combinative scheme.

For each of these approaches, character is little more than an enabler of action. The mental states, explicit or implicit, that motivate agentive characters to plan and act—the fears or hopes for possible futures—are not a part of the equation. Some structuralist thinkers took a character-centric approach, though. Bremond [1980] saw plot as a “network of possibilities” that underlies each action, with characters thinking several moves ahead in a branching model not unlike game theory. In a Propp-style study of French folk-tales [Bremond, 1970], he finds a pattern in which characters cycle between positive and negative affect states.

Structuralism fell out of favor with Derrida and other proponents of deconstruction. Barthes himself turned away from the search for a universal symbolic underpinning to narrative in his later works, advocating instead for a reader-oriented approach to narrative meaning [Barthes, 1978]. Although the symbols we reviewed in our discussion of cognitive psychology are also subjective, in that each reader builds his or her own situation model, deconstructionists are far more cagey about proposing universal symbols and relations that can apply to many texts. In *S/Z* [Barthes, 1974], Barthes proposes a set of “codes” that function as threads in an interwoven search for meaning on the reader’s part. The hermeneutic code, for instance, gives a sense of suspense to the reader by withholding answers to important questions or delaying an expected event. The five codes apprehend the text as a plurality, a “galaxy of signifiers, not a structure of signifieds,” that “has no beginning; it is reversible; we gain access to it by several entrances, none of which can be authoritatively declared to be the main one.” The reader is an active participant in constructing threads of narrative meaning, rather than a passive receiver of a single authoritative model of the text.

Although our approach is structuralist in character, we do not make a claim that a structuralist approach to literature is inherently superior to such a receiver-oriented approach for the task of designing a formal representation of story meaning. Rather, we pragmatically find that in the absence of a wide-coverage understanding and inference engine that can read a text and simulate such processes of reading as anticipating future actions and detecting questions needing answers, we cannot systematically compare stories by beginning with a process like the hermeneutic code and working backwards to the textual facets that trigger a response. We must instead start with connections we can find at the discourse, textbase and/or *fabula* level. At that point, if we are studying affect, we can move forward to predicting a normative reader response to a story based on responses to similar stories. However, the lesson that we can draw from post-structuralism is that subjective differences between receivers, as well as plural readings by a single receiver, cannot be marginalized in a symbolic model of narrative discourse. As we saw in Chapter 1, stories are told to communicate information and to evoke an affectual response from the receiver. We re-read books and re-watch films not to be reminded of the plot, but to repeat the experience of being in the narrative world. Unlike most annotation tasks, where inter-annotator agreement is unequivocally bad, a narrative interpretation task should elicit, embrace and study
the differences between subjective encodings of the same text.

Since the decline of structuralism, literary theorists who claim that there are absolute universals of any kind inherent across the narrative mode of discourse have been few and far between. Even the notion of such a goal is not often conceded by academic theorists (which is why we have turned to psychology for the foundations of a new representation). One recent trend, though, is to apply the **theory of mind** to literature and the semiotics of narrative. This approach takes as its unit of analysis the presence of a conscious entity in the story, capable of identity, self, subjectivity and experientiality [Palmer, 2007]. The receiver must connect to conscious agents within the narrative who are perceiving and transmitting the story-world information (roughly analogous to the narrating agents between the layers of Bal’s *fabula-story* distinction). Reading a novel is akin to following the thought processes of these agents. Palmer [2010] applies this mode of analysis to Dickens's *Little Dorrit*, which we included in our study in Chapter 2, and finds a social interplay among the characters as they signal, withhold, and yield information to one another. They derive power and status from this *intermental* (collective) thought, as opposed to individual or private thought.

The foundation for this reading is **attribution theory**, an area of cognitive psychology which (among other purposes) describes how observers attribute states of mind to other conscious entities—not only goals and plans, as we reviewed, but emotions and the entities’ own attributions regarding the minds of still other entities [Jones and Nisbett, 1972]. In other words, what is important is not just what is known, and what is known to be unknown, but what is known to each character, what each character knows about what is known to each other character, and so on, in what we will represent as a system of nested **agency frames**. (The outermost or “bottom” frame represents what is known to an agent; the second frame from the bottom represents what an agent knows about what other agents know, and so on. All stacks rest on an objective foundation that we will call “ground truth.”) This model is similar to that of the *private state frame*, which Wiebe et al. [2005] devised as a fine-grained annotation scheme to assist in the automatic identification and extraction of opinions, emotions, and sentiments in text. The ability to “mind-read” is also known to be an important milestone in development; autistic children have abnormal abilities to see the world through the eyes of others [Baron-Cohen, 1989].

As attribution is key to understanding stories as simple as fairy tales, nested agency frames should be expressible in our new narrative representation. In “Little Red Riding Hood”, we know that the wolf is masquerading as the girl’s grandmother, and we know that the wolf knows that fact, and we know that Little Red Riding Hood does not know that her grandmother is really the wolf, and we know that the wolf’s goal is to maintain such a belief on the part of the girl [Chen and Fahlman, 2008]. In some tellings, we suspect that the girl is transitioning to a new state of understanding during the dialogue in which she expresses continuous surprise at her supposed grandmother’s wolf-like features (“What big, sharp teeth you have!”). In both “The Wily Lion” and “The Fox and the Crow”, the predator’s plan is to instill a belief in his prey that the predator thinks the prey is almost appealing in some way. In both cases, the deception hinges on the difference between the reality of the predator’s thoughts, known to the receiver and to the predator, and the feigned reality purported to the prey. To understand these stories is to constantly separate not only what actions happen before other actions, but also what mental states are occurring and
who is associated with them. As Zunshine [2006] argues in her theory-of-mind reading of Woolf’s *Mrs. Dalloway* and other modernist works of literature, the narrator’s “engagement of our metarepresentational capacity”—its method of leaving clues with which we can read the minds of conscious characters—is a universal that gives the novel its currently familiar shape.

### 3.2.3 Implemented Understanding: Scripts, Plans and Plot Units

Those who have implemented story understanding systems have typically come from the artificial intelligence tradition, though they sometimes work in concert with cognitive psychologists to arrive at a cross-domain theory of the way the mind represents narratives (e.g., [Charniak, 1972; Schank and Abelson, 1977]). It was one of the first problems considered by what we now call natural language processing: The phrase “story understanding” has often entailed a system that, given a textual story, can answer yes/no and “why” questions that demonstrate inference, retell the story, or generate a summary by choosing key aspects and reformulating them into new sentences [Lehnert *et al.*, 1983; Reeves, 1991]. Unfortunately, as we mentioned earlier, such a system needs a large and laborious modeling of world knowledge in order to interpret natural language input to the point of inferring causes, consequences, ethics, and key points (that is, comprehensive semantics). The previously discussed work in psychology has informed us that human story comprehension is a fusion between the propositions in the textbase and the world knowledge of the receiver—but models of comprehension such as causal networks only record *that* an inference is made by a human reader, without describing *how*. To build a story understanding system, those methods of inference must be concretized, a problem that has yet to be solved at scale [Lehnert, 1994].

The early researchers attempted to “start small,” restricting their input to a very limited set of stories that each invoke only a slice of domain knowledge, and then broaden coverage over time. The broadening never quite arrived, though, and when the tide shifted to corpus-based methods in the 1990s, narrative as a discourse mode of study fell by the wayside. While a few long-term projects are taking steps toward deep understanding with broad coverage [Mueller, 2006; Zarri, 2010], work in the computational modeling of narratives has been performed periodically in different areas—sometimes from a discourse context, sometimes from an agent-centric context, sometimes from a generation context, sometimes from a ludic (interactive gameplay) context. Each context carries with it a particular set of constraints on the narrative representation. Story generation and interactive narrative, for instance, require prescriptive models, with rules for story structure and character behavior defined precisely enough to prevent nonsensical output. A descriptive model such as ours trades off precision in order to be robust enough for a data-driven (corpus-based) analysis. In this section, we review some of the story representations from both categories that have been devised to build narrative-savvy systems.

The work in understanding at Yale in the 1970s [Schank and Riesbeck, 1981] saw narrative as a way of controlling the combinatorial “explosion” of inference that occurs when reading discourse. This is because the scope of possible logical interpretations of two sentences that have a functional connection in a model of discourse is smaller than the scope of
possible interpretations of both sentences as seen individually. “Narrative,” then, is a mode of discourse where agents have goals and pursue those same goals in a logically consistent pattern, and a library of narrative facets can guide a system as it interprets the meaning behind a textual story such as a news article.

The common ground between artificial intelligence and the psychology of narrative comprehension was seen early on in work describing the inferences involved in recognizing the goals and plans of agents based on their actions [Schmidt et al., 1978]. The efforts to formalize a system of goals in commonsense psychology continue today, in particular by Hobbs and Gordon [2005; 2010]. They describe logical axioms for the process of identifying goals, developing plans for achieving them, monitoring the outcome of those plans, and modifying those plans as needed. For example, if an agent wants $e_1$ and believes $e_2$ causes $e_1$, that desire will cause the agent to also want $e_2$ (as a subgoal). This recalls the models of Trabasso and Graesser, but adds a notion of importance in which agents give their goals a partial ordering by preference. A rational agent, then, is one that first pursues those goals that are most important. Many stories hinge on this notion of weighing one goal against another, the essence of the dilemma. A protagonist must decide whether to go for the job promotion or fulfill a family obligation, or whether the pursuit of love is worth the alienation of one’s tribe.

The script was introduced by Schank and Abelson [Schank and Abelson, 1977; Cullingford, 1981] as a frame-like structure [Minsky, 1975] that stores procedural knowledge about some process: what happens, and in what order. In the canonical example, the RESTAURANT script gives a sequence of “scenes” (entering, ordering, eating, exiting), each of which containing first-order actions (customer enters restaurant, customer sees tables, customer sits). In the SAM (Script Applier Mechanism) system, when the restaurant script is activated by an action such as “John entered the restaurant,” the script triggers expectations for what may happen next. Though scenes are an attractive organizing principle for narratives, such hard-coded scripts fell out of favor due to their rigidity. Contemporary analogies to scripts, such as narrative event chains [Chambers and Jurański, 2008a] and certain aspects of the FrameNet project [Johnson and Fillmore, 2000], use corpus-driven statistical models to generate or evaluate scripts.

The plan was devised by Schank’s students as a knowledge structure that was focused around agents rather than episodes. Plan formalisms describe the motivations and ultimate goals for actions that might appear in the text. They can be combined and modified more flexibly than scripts. Wilensky’s PAM (Plan Applier Mechanism) system [Wilensky, 1978a] understands textual stories by inferring the intentions of the story’s characters from their actions. Unlike scripts or Trabasso’s causal networks, PAM could represent opposing goals as well as chains of why questions [Wilensky, 1978b]. It might infer that John asked Mary where he could find a restaurant because John was hungry, but that John’s hunger does not need to be explained, because people normally need food. PAM was able to stitch together stories based on chains of intersecting plans, and had a library of both plan architectures and goal transformations. For instance, it understood goal subsumption as form of long-term planning in which an action is designed to address a potentially recurring goal rather than one which is clear and present. From this principle, it further asserts that marriage is a plan that prevents recurring episodes of loneliness which may occur later in life. While PAM
suffers many of the same issues with rigidity that SAM does, it introduces an attractive architecture for modeling the way in which plans from different agents intersect, whether in competition, cooperation, motivation, subsumption, or another relationship.

The formal treatment of goals, plans and beliefs evolved into an action control architecture called BDI (beliefs, desires and intentions) [Bratman, 1987; Rao and Georgeff, 1995; Busetta et al., 2003; Konolige and Pollack, 1989]. BDI models not only what is known and unknown to a certain agent, but what the agent’s goals are and what actions are possible for it to take that might eventually satisfy those goals. In one form, it defines formalisms for the theory-of-mind question of what agents know about the world and about what other agents know [Rapaport, 1986]. BDI has matured over the years to the point where it can approximate the process of rational decision-making and plan-making using limited evidence, emphasizing the subtleties of intention behind each action [Cohen and Levesque, 1990; Pollack, 1990]. Although BDI was not designed as a representation for narrative discourse, its agent-centric approach has inspired similar architectures that allow interactive narrative systems to have plausible, value-driven characters [Peinado et al., 2008; Damiano and Lombardo, 2009]. Similarly, the partial-order planning architecture, sometimes combined with features of BDI, is widely used in narrative generation and interactive-narrative systems [Mateas and Stern, 2003; Riedl and Young, 2004; Mott and Lester, 2006; Winegarden and Young, 2006; Barber and Kudenko, 2008]. Planning is also a useful approach to generating surface-level narrative in both textual discourse and visual media [Callaway and Lester, 2002; Jhala, 2004]. As was the case with PAM and SAM, these systems trade off robustness and expressibility in order to arrive at a level of formal precision sufficient for controlling actions and describing narratives. Their scope is limited to just those narratives consisting of actions and goals that have been modeled by hand (though some BDI systems include a high-level language to allow domain experts to extend the knowledge base as needed [Rao et al., 1992]).

One recent project has aimed for a more abstract, yet computable representation of actions, mental states and agent behavior. SCON [Chen and Fahlman, 2008] uses a semantic network representation in which nodes represent entities such as physical objects, types of actions, and actual actions that have occurred at some point in time. The system represents actions in a frame-slot representation [Minsky, 1975] but does not understand their consequences in the sense of Mueller’s first-order representation [Mueller, 2004], in that SCON cannot infer implied information or answer general questions about a story. What it does understand deeply is the epistemological component: what characters know, and what they know about what others know. Each action is situated in a mental context tied to the time in which the action occurred. The mental context is agent-specific. In their encoding of the Brothers Grimm’s telling of “Little Red-Cap” (a.k.a. “Little Red Riding Hood”), the wolf’s intention to eat the girl is different from the girl’s belief in the wolf’s intention, because the girl believes that the wolf is her grandmother. Their system can detect contradictions within mental states and answer simple questions about what characters know. However, it does not currently model goals and plans, so there is no way to understand the crucial point that the girl’s belief in the wolf’s identity is itself the wolf’s intention and part of a larger plan. Our project is similar in approach to SCON, but it provides a more expressive symbolic vocabulary. 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 and Pacuit, 2008; Löwe et al., 2009; Nissan, 2008]; we see this as a promising approach.

Plot Units

We now conclude our tour of prior models of narrative discourse with a surprisingly versatile formalism called plot units, developed as part of a semantic story understanding system called BORIS [Lehnert et al., 1983]. BORIS was capable of parsing, interpreting and answering questions about simple stories in a small knowledge domain. One of its functions was to summarize the key points of a narrative; to do this in a thematically aware manner, Lehnert [1981; 1984] devised a data structure that represents the affectual state of each character. As BORIS reads the story, it creates a linear map of temporally ordered affect states for each agent. There are three possible states:

• + (Positive Event): Events that please the agent in question
• – (Negative Event): Events that displease the agent in question
• M (Mental State): Mental state of the agent in question

The affect states are then connected by instances of four types of directed arcs:

• Motivation (m) always points to a mental state M, and from the state which caused M
• Actualization (a) always points from a mental state M, and to a + or – state intended by M
• Termination (t) points from one mental state or event to another that the first displaces (replacing one goal with another, or having the positive affect of a + displace the negative affect of a –)
• Equivalence (e) points from one mental state or event to an identical copy (useful for when two agents perceive the same event in different ways)

In all, there are 15 legal pairwise configurations that function as “lexemes” of plot, as seen from an agent-affect perspective. All arcs and nodes are instantiated in a domain associated with an agent. Some arcs may stay in the same domain (belonging to the same agent) or cross domains (describing a particular causal effect that one agent has on another). Each configuration corresponds to a thematically interesting narrative event of some kind. The notion of a “mixed blessing,” for instance, has not been modeled by any previous representation we have examined, yet is expressed with a single (e) arc from a negative state to a positive state in the same agent domain (indicating that an event is at once good and bad for the agent). Figure 3.4 depicts the 15 basic units along with their thematic interpretations.
The greatest virtue of Lehnert’s model is that like Trabasso’s causal arcs, simple plot units can be chained together to form compounded units of arbitrary length and complexity. The expressive range is more powerful than that of the GRTN, though, due to the separation of events and states into agent-oriented domains. One event can have multiple consequences to different agents, occurring in parallel from different perspectives. There no longer needs to be one central protagonist. Lehnert identified some 30 complex plot units describing exchanges between two characters such as the double-cross (where one agent requests an action and instills a mental state in another, only to have the latter agent trigger an event that helps himself and hurts the requester).

“The Wily Lion” can be reduced to a plot-unit representation by virtue of the fact that it hinges on a request that purports to have a positive affectual impact on both characters, but actually has a positive impact for the requester and a negative affect for the requestee. A plot-unit representation of this fable chains five simple plot units to achieve this effect. Figure 3.5 shows the plot-unit mapping in which the fox is motivated to eat, fails, considers his problem, and has success in flattering the bull, who believes he is enabled to make himself more handsome. The bull succeeds in improving his image, but the same event allows the lion to resolve his hunger, which causes a grave loss on the bull’s part but is a success to the lion.

Lehnert’s model scores highly against our criteria. It is robust, in that it does not depend on a full semantic understanding of the story per se (although it was originally conceived in conjunction with such a system). The chainable nature of plot units makes them a more expressive formalism than a grammar or RST-style representation, allowing parallel causal chains, multiple agent perspectives, and so on. Plot units support implied information, rather than being a rearrangement of the textbase alone. The semantics of the node and edge types are well-chosen to be able to emulate subtle and complex narrative exchanges. Comparisons between stories can be made using complex plot units as a mediator (both “The Wily Lion” and “The Fox and the Crow” might satisfy a betrayal plot unit). Our
representation strives to replicate each of these advantages.

The plot-unit model, though, does not capture other elements that are useful in a descriptive representation. There are very few time operators, for instance, and it is not obvious that every event in the story that is relevant to the plot must necessarily have a negative or a positive affect to an agent. Further, a ternary \(+/-/M\) system is somewhat coarse. \(+/-\) can refer to either events or states, and \(M\) does not address the epistemic question (what one agent knows or believes about another’s knowledge frame). The distinction between hypothetical events (such as goal states) and actual events is unclear; it is impossible, for example, to indicate that a multi-step plan was completely abandoned by a character because another character brought about the ultimate goal on his own accord. The lack of hypotheticals also makes goal hierarchies awkward. In Figure 3.5, it is not quite right to say that “is very hungry” presents a problem by motivating the lion to want the bull to remove its horns; it is more of a restatement of the same problem that previously
failed when the bull became fearful. These issues are partly due to scope, in that BORIS contained elements that worked orthogonally to plot units and handled aspects such as time. Our new representation, divorced from a semantic inference engine, can extend plot units along these lines.

Ryan [1991; 2007] has proposed a representation called a “recursive graph model” that inherits aspects of Lehnert’s architecture (including an open graph structure and character-specific domains), but adds a much richer set of primitives for describing mental states. Each agent has a set of five distinct domains: K-worlds, which are epistemic (beliefs, projections and retrospections), O-worlds, which are private or social obligations, W-worlds, which are desires and fears, G-worlds, which are active goals, and P-worlds, which contain the plans through which characters seek to fulfill their active goals. Physical events, in a timeline, are separated from mental events, which are grouped by agent. To our knowledge, though, Ryan’s model has yet to be implemented in an annotation interface or corpus collection project.

Plot units have also been influential in statistical approaches to understanding textual narrative. One recent effort takes a large step toward extracting plot units from unstructured text by classifying the affectual states implied by various clauses with respect to agents that appear as named entities [Goyal et al., 2010a]. Nackoul [2010] develops a natural-language template for describing plot units as well as a system that uses these templates to search for plot units as they appear in narratives as diverse as Macbeth and legal case briefs. Similarly, Appling and Riedl [2009] return to Lehnert’s original intention for plot units, summarization, with a system that uses conditional random fields to label affectual states, events and relational links as they appear in surface text.

3.2.4 Conclusion

The preceding literature review has featured symbolic models of narrative discourse from a variety of fields and intents. In aggregate, they present a set of tradeoffs: between formality and robustness, between an event-centric and agent-centric view of what a story is, between prescriptive and descriptive. The most promising aspects of each model have guided the design of our own contribution. In particular, we are motivated to include a system which can express goals, subgoals, plans, beliefs, attempts to achieve goals, and goal outcomes, as the studies from cognitive psychology strongly suggest that these are hard-wired into the human narrative instinct. In terms of structure, we are attracted to the notion of a small but highly recombinable lexicon of nodes and arcs, as we saw in Trabasso’s and Lehnert’s models. These are not only abstract and consistent from story to story, enabling contrastive studies; they are also accessible and can be tractably extracted from surface text by automatic taggers. The theory of mind offers a favorable template for modeling complex epistemic interactions in the context of separate agents; these interactions are often behind the thematic crux of a story. Finally, linguistic work along the lines of Propp, Polanyi, Labov and Waletzky and Passonneau tells us that we can find semantic similarities between stories in a corpus without committing to a complete semantic understanding of narrative fabula. In the next section, we describe Story Intention Graphs, which attempt to synthesize these insights in a new representation for story annotation, reasoning and comparison.
3.3 Story Intention Graphs

Narrative is an interplay between the minds of agents, the actions they take, the events which befall them, and the perception and transmission of that content in a communicative artifact. In this section, we propose a representation of a story that reifies these as nodes and arcs (relations) in a semantic network. We call the schemata itself the “Story Intention Graph” (SIG), and each instance of story annotation using this model a “SIG encoding.” The SIG is a constructionist model, in that it brings out coherence at both the local and global levels: what events happen, when, why, and to whom. Like the previous models we examined, in a SIG the entire discourse is modeled in a single, integrated data structure. It is descriptive, rather than prescriptive.

A SIG consists of three interconnected subgraphs called layers:

**Textual layer.** Analogous to the *text* layer in the van Dijk and Kintsch [1983] model, or the *discours* to Todorov, this is a linear vector of nodes that contain the *utterances of the original discourse* that is being modeled. While we only deal with textual discourse in the present study, nodes in the textual layer can also represent snippets of other kinds of media, such as oral storytelling. Each node contains anywhere from one proposition’s worth of text to a paragraph or a passage, depending on its role in the overall structure of connected relations. Collectively, all these nodes represent the story as it is told from the telling’s start to the telling’s finish.

**Timeline layer.** Nodes in the timeline layer formally encode *story-world happenings* that have been expressed in the textual layer, such as events and statives. These nodes are arranged in a timeline that represents the sequence of story-world time. This layer is analogous to Todorov’s *histoire*, van Dijk and Kintsch’s *textbase* and the Formalist *fabula*. It represents the stated story content from the beginning of the story’s chronology to the end of the story’s chronology. Each node is annotated with the identity of the agent, if any, that is responsible for the narrated happenings (e.g., the perpetrator of an event). A more complete knowledge representation of the content in question, such as a predicate-argument structure, may also be attached to each node, though none is required.

**Interpretative layer.** The interpretative layer is analogous to the cognitive situation model. Here, nodes represent *goals, plans, beliefs, affectual impacts, and the underlying intentions of characters (agents)* as interpreted by the story’s *receiver*. This includes both content that is directly stated (duplicating timeline-layer content) and content that is implied, but never stated outright, in the story as it is narrated. Its purpose is to relate *timeline-layer and textual-layer content by their motivational, intentional and affectual connections*, as opposed to their temporal connections as in the other two layers. For example, five actions in the story can all be intentional attempts to reach the same implied goal, which is represented as a node in the interpretative layer even though the narrator never explicates it. Collectively, the interpretative layer represents a receiver’s agent-oriented interpretation of the narrative, with connections back to the stated content (in the textual and timeline layers) that justifies it.
This section introduces the set of node types and relations (arc types) that constitute the three layers of the SIG. Table 3.3 gives a summary of the node and arc types we will describe. To illustrate the instantiation of the schemata for encoding a particular story, we will apply the SIG to “The Wily Lion” and compare the result to those of previous models.

This approach differs from prior work in two important areas:

1. The discourse order of the surface text is preserved alongside the chronological order of the narrated content. That is, the SIG includes two temporal orderings of the stated story content, rather than one: an encoding of the discourse fragments in which they appeared in the narrated discourse (telling time, in the textual layer) and an encoding of the same content in the chronological order of the story-world being described (story time, in the timeline layer). The previous models we examined disregarded either the telling time (GRTNs, plot units) or the story time (grammars). In the SIG, both orderings are present and cross-referenced. In the next section, we will see how this is useful for modeling narrative discourse.

2. Previous models conflate what we call timeline and thematic content. For instance, plot units are built on a +/−/M system, with the + and − indicating both an event and its affectual impact on an agent. GRTNs categorize story actions among a set of mutually exclusive, goal-oriented labels (Goal, Action, Outcome and so on). Story grammars include both temporal and goal-oriented rewrite rules (EPISODE and GOAL_PATH, respectively). These conflations make certain narrative scenarios difficult or impossible to describe, such as the hidden agenda, where one action serves distinct purposes in two separate plans. In the SIG, the timeline layer encodes only the temporal organization of the textbase, with its discrete nodes of narrated story-world content. Goal-oriented labels appear in the interpretative layer, which, while separate, is cross-referenced to the timeline and textual layers. We demonstrate in Appendix B that this modular approach enables the schemata to have a wide expressive range for describing many types of narrative situations, including hidden agendas.

### 3.3.1 Textual and Timeline Layers

The textual and timeline layers of the SIG include nodes for the surface form of the discourse and for the textbase form of the narrated story-world. The node and arc types are chosen to organize the textbase into a semantic network based around time.

In the textual layer, the discourse is divided up into fragments (continuous spans). Each fragment is represented by a Text node. Text nodes are chained together by followed by arcs so that the order of nodes in the chain reflects the order in which the fragments appear in the original discourse. The textual layer, then, encodes the “telling time” of the story. Each Text node is linked to a node in the timeline layer that represents an equivalent textbase happening. Figure 3.6 illustrates the beginnings of a SIG encoding for “The Wily Lion”, including three text fragments in the textual layer and their equivalents in the timeline layer. The interpreted as arc indicates equivalence. For instance, the first Text node in the chain of Text nodes, TE1, represents the first sentence of the discourse, “A Lion watched a fat Bull feeding in a meadow.” This node is attached with interpreted as to a Proposition node.
CHAPTER 3. STORY INTENTION GRAPHS

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Name</th>
<th>Usage/Signified Element</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TEXTUAL LAYER</td>
<td></td>
</tr>
<tr>
<td>TE</td>
<td>Text</td>
<td>Continuous span of surface discourse</td>
</tr>
<tr>
<td>f</td>
<td>Follows</td>
<td>Ordering of text spans in a discourse</td>
</tr>
<tr>
<td></td>
<td>TIMELINE LAYER</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>State</td>
<td>An instant of story-world time</td>
</tr>
<tr>
<td>P</td>
<td>Proposition</td>
<td>A unit of discrete story-world content, such as an occurring action or event, typically pertaining to an agent. Can include a more complete knowledge representation of the narrated happening</td>
</tr>
<tr>
<td>T</td>
<td>Timeline</td>
<td>A continuum of time states in the story-world in a single modality</td>
</tr>
<tr>
<td>f</td>
<td>Follows</td>
<td>Ordering of States in a Timeline</td>
</tr>
<tr>
<td>ia</td>
<td>Interpreted as</td>
<td>Equivalence between TE and P nodes</td>
</tr>
<tr>
<td>i</td>
<td>In</td>
<td>Connects a State to its Timeline</td>
</tr>
<tr>
<td>ba</td>
<td>Begins at</td>
<td>Connects a Proposition to its temporal initiation State</td>
</tr>
<tr>
<td>ea</td>
<td>Ends at</td>
<td>Connects a Proposition to its temporal termination State</td>
</tr>
<tr>
<td>r</td>
<td>Referenced by</td>
<td>Connects a Timeline to a P or I node that incorporates it modally</td>
</tr>
<tr>
<td>e</td>
<td>Equivalent</td>
<td>Connects State nodes referring to the same moment in two Timelines</td>
</tr>
<tr>
<td></td>
<td>INTERPRETATIVE LAYER</td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>Interpretative Proposition</td>
<td>A unit of story content, equivalent to a P node in the Interpretative space. Either Hypothetical (H), Actualized (A) or Prevented/Ceased (PC) with respect to each State of the main Timeline</td>
</tr>
<tr>
<td>G</td>
<td>Goal</td>
<td>Indicates that certain I, G or B nodes are the goal of an agent</td>
</tr>
<tr>
<td>B</td>
<td>Belief</td>
<td>Indicates that certain I, G or B nodes are the belief of an agent</td>
</tr>
<tr>
<td>A</td>
<td>Affect</td>
<td>The baseline affectual state of an agent</td>
</tr>
<tr>
<td>i</td>
<td>In</td>
<td>Connects an I, G or B node to the G or B frame in which it is situated</td>
</tr>
<tr>
<td>ia</td>
<td>Interpreted as</td>
<td>Equivalence between a P node and an I, G or B node</td>
</tr>
<tr>
<td>im</td>
<td>Implies</td>
<td>Implication by a P node of an I, G or B node</td>
</tr>
<tr>
<td>a</td>
<td>Actualizes</td>
<td>Links a P node to an I, G or B node when the reader infers that the latter becomes actualized because of the former</td>
</tr>
<tr>
<td>c</td>
<td>Ceases</td>
<td>Links a P node to an I, G or B node when the reader infers that the latter becomes prevented/ceased because of the former</td>
</tr>
<tr>
<td>wc</td>
<td>Would cause</td>
<td>Link between one I, G or B node and another that is sufficient for its actualization</td>
</tr>
<tr>
<td>wp</td>
<td>Would prevent</td>
<td>Link between one I, G or B node and another that is sufficient for its prevention/cessation</td>
</tr>
<tr>
<td>pf</td>
<td>Precondition for</td>
<td>Link between one I, G or B node and another that is necessary for its actualization</td>
</tr>
<tr>
<td>pa</td>
<td>Precondition against</td>
<td>Link between one I, G or B node and another that is necessary for its prevention/cessation</td>
</tr>
<tr>
<td>ac</td>
<td>Attempt to cause</td>
<td>Indicates intention by the agent of a P node to actualize an I, G or B node</td>
</tr>
<tr>
<td>ap</td>
<td>Attempt to prevent</td>
<td>Indicates intention by the agent of a P node to prevent/cease an I, G or B node</td>
</tr>
<tr>
<td>p</td>
<td>Provides for</td>
<td>A positive affectual impact of an I, G or B node (traversing to A)</td>
</tr>
<tr>
<td>d</td>
<td>Damages</td>
<td>A negative affectual impact of an I, G or B node (traversing to A)</td>
</tr>
</tbody>
</table>

Table 3.3: Summary of the types of nodes and arcs that constitute Story Intention Graphs. Node types have capitalized symbols; arc types have lowercase symbols.
Figure 3.6: Fragment of a SIG encoding showing textual-layer nodes, as well as Proposition nodes in the timeline layer. A non-contiguous subset of “The Wily Lion” is encoded.

containing an equivalent textbase unit. In this case, the unit is labeled with a propositional equivalent, \texttt{watch(lion, feed(bull, meadow))}.

The size of each fragment of surface text is set so that the resulting Text node can be connected to a uniformly equivalent P node in the timeline layer, which in turn is connected to an expression of agentive intent in the interpretative layer. In a concise discourse such as a fable, this is typically of clause or sentence length; in other cases, a longer passage may be reducible to a single functional unit with respect to an agent-oriented reading of the narrative.

Figure 3.6 shows the temporal structure in the textual layer with \textit{followed by} arcs. In a more complete SIG encoding, the Proposition nodes are also temporally ordered (hence the name of the layer); crucially, though, they are arranged in a timeline that corresponds to “story time,” the chronology of the story-world. The question of how to structure this arrangement is non-trivial. Time is the most fundamental discourse relation in a story, and most thematic content depends on an ordering of events (for instance, actions are followed by their consequences). A depiction of temporal ordering is common to all the representations we examined in Section 3.2, but each model simplified to some degree the many temporal relationships that can be found within a story. The relationship between story time and telling time can be quite complicated (see Mani et al. [2005] for a comprehensive review). Events in a story can occur over long or short periods of time, overlap, terminate one another, refer back or forward to other points in time, and cross over into hypothetical or imagined modalities. Even given a representation of time, the process of parsing the tense and aspect of narrative rhetoric (whether in English or another language) into a
formal understanding of time is quite complex, subject to decades of work in linguistics
(e.g., [Comrie, 1976; Comrie, 1985; Halliday, 1976; Nerbonne, 1986; Hornstein, 1990; Vlach, 1993]), natural language processing [Hinrichs, 1987; Webber, 1987; Passonneau, 1988; Mani and Pustejovsky, 2004], and artificial intelligence/database theory [Allen, 1991; Özsoyoglu and Snodgrass, 1995; Terenziani and Snodgrass, 2004]. Temporal and modal relations have also been singled out as bases for a discourse annotation scheme, TimeML [Pustejovsky et al., 2003a], and associated annotated corpus, TimeBank [Pustejovsky et al., 2003b]. While a detailed inquiry into the process of understanding or representing time is beyond the scope of this thesis, we will later revisit the relationship between a formal representation of time and English tense and aspect (Section 4.4).

We propose here a representation of time for the SIG that is robust, computable and amenable to manual annotation: The structure of the timeline is based on event *intervals* in the tradition of Allen’s classic work on temporal reasoning [Allen and Ferguson, 1994]. Each P node takes place over an interval, which is a pair of states (points on a linear timeline). The states have a complete ordering which we can express with an enumeration, \( t_1 \ldots t_n \).

In sum, there are four types of nodes in the textual and timeline layers:

**Text node (TE)** Represents a continuous span of surface discourse corresponding to a textbase happening (P) node and agentive interpretation in the interpretative layer.

**State node (S).** Represents an instant in time in the story-world. Each State has an associated time index \( t \), a natural number:

\[
\text{t(S)} \in \mathbb{N}
\]

**Timeline node (T).** Represent a continuum of states in the story-world. There must be at least one main Timeline node, dubbed the **Reality timeline**, in a SIG encoding. As we will see, modal situations such as imagined past events are expressible with additional Timeline nodes.

**Proposition node (P).** An encoding of story-world content (a textbase happening, such as an event) that corresponds to a span of surface discourse. The happening occurs at a single State or an interval between States. (The interval can be unbounded in the case of events that never end or whose ending points are unimportant.) Epistemically, P node can belong to any **modality**: The content can depict an occurrence in the story-world’s reality, an imagined concept such as a fear, an opinion or a metatextual comment. (The modality is set by the Timeline node associated with the Proposition node through the arcs we describe below.) If the story content involves an intentional agent, that agent is associated with the node as metadata.

Throughout this chapter, we illustrate P nodes with propositional (predicate-argument) encodings; however, despite the node’s name, the type of encoding used to represent textbase content within a P node is unimportant to the SIG schemata. A P node may, for instance, have no encoding whatsoever—in this case, it only marks that a certain span of story text occurs at a certain story-world time, in a certain modality, featuring a certain discrete agent. In Chapter 5, we construct some SIG encodings using
this “placeholder” technique, and others where each P node features a constructed propositional equivalent of the text span associated with the corresponding Text node.

There are seven relations in these layers, the first five of which are:

**Followed by (f).** Placed between one Text node and the Text node that immediately follows the first node in the original discourse. Also traverses between a State node and State node of the same timeline that immediately follows it. Logically, $f$ is transitive, although the implied arcs are not drawn in the SIG:

\[
\begin{align*}
  f(S_1, S_2) \land f(S_2, S_3) & \Rightarrow f(S_1, S_3) \\
  f(S_1, S_2) & \Leftrightarrow t(S_1) < t(S_2)
\end{align*}
\]  

**Interpreted as (ia).** Traverses between a Text node and a Proposition node which represents the textbase equivalent of the discourse fragment associated with the Text node. There is a many-to-many relationship permitted with ia: a Text node can invoke several Proposition nodes, and each Proposition node can be justified by several discourse spans.

**Begins at (ba).** Traverses between a P node and the State at which the proposition first takes effect in the story-world. Such a relationship is not necessary for every P node. Propositions that do not link to a State with $ba$ do not have a start time that can be inferred from the content of the corresponding Text node.

**Ends at (ea).** Traverses between a P node and the State at which the proposition culminates, stops or ends. Such a relationship is also not necessary for every P node. Propositions that do not link to a State with $ea$ do not have an ending time that can be inferred from the content of the corresponding Text node.

If both $ba$ and $ea$ are given for a Proposition node, the beginning state must be followed by the ending state:

\[ ba(P, S_1) \land ea(P, S_2) \Rightarrow f(S_1, S_2) \]

**In (in).** Traverses between a State node and the Timeline node representing the scope of story-world time in which it exists. A Timeline is said to “contain” all of the States that link to it with in, as well as all of the Proposition nodes that link to those States with $ba$ or $ea$. We use the $\in$ notation as shorthand for contained by:

\[
\begin{align*}
  in(S, T) & \Leftrightarrow S \in T \\
  (ba(P, S) \lor ea(P, S)) \land in(S, T) & \Leftrightarrow P \in T
\end{align*}
\]  

Note that a Proposition can only belong to one timeline:

\[
\begin{align*}
  ba(P, S_1) \land ea(P, S_2) \land in(S_1, T) & \Rightarrow in(S_2, T) \\
  ba(P, S_1) \land ea(P, S_2) \land in(S_2, T) & \Rightarrow in(S_1, T)
\end{align*}
\]
The union of a Timeline node and all of the State nodes and Proposition nodes that it contains through in, ba and ea is collectively known as a timeline.

Figure 3.7 shows a representative example of these two layers of the SIG. Three fragments of “The Wily Lion” are mapped onto four propositions, including one modifier (one Text node has two outgoing interpreted as arcs). The temporal relationship between the first two events is what Allan terms meets: one’s start time is the other’s end time. At State S₄, a stative representing the bull’s identity as the lion’s prey begins; this is in essence a “become” action (the bull becomes the lion’s prey). In all, the four propositions involve four time states, each of which is seen to be in the Reality timeline.

A complete textual- and timeline-layer encoding of “The Wily Lion” is shown in Table 3.4. The first two columns are the vector of Text nodes in the textual layer; the followed by arcs that join adjacent State and Text nodes are not shown. Also implied but not shown are the interpreted as arcs traversing from each Text node to the associated Proposition node(s) on its respective row. The outgoing arcs incident to each Proposition node are shown in the rightmost column—begins at and ends at arcs traversing to State nodes.

**Story Time vs. Telling Time**

Since the SIG features a mapping between the discourse ordering of events and the story-world ordering of events, it allows us to study a discourse in terms of the ordering and pacing of its fragments with respect to the story-world being described. From the nodes and arcs we have introduced, we can draw a “plot” of the discourse in which the horizontal axis is
<table>
<thead>
<tr>
<th>Node</th>
<th>Text Node content</th>
<th>Arcs</th>
</tr>
</thead>
<tbody>
<tr>
<td>$TE_1$</td>
<td>A Lion watched a fat Bull feeding in a meadow</td>
<td>$P_1$ watch(lion, feed(bull, meadow)) ba(S$_1$); ea(S$_5$)</td>
</tr>
<tr>
<td>$TE_2$</td>
<td>and his mouth watered when he thought of the royal feast he would make</td>
<td>$P_2$ thought(lion(potentialFuture( identity(bull, feast))))) ba(S$<em>2$); ea(S$</em>{13}$)</td>
</tr>
<tr>
<td>$P_3$ watered(mouth(lion))) ba(S$_2$); ea(S$_3$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$TE_3$</td>
<td>but he did not dare to attack him,</td>
<td>$P_4$ ~attack(lion, bull) ba(S$<em>5$); ea(S$</em>{13}$)</td>
</tr>
<tr>
<td>$TE_4$</td>
<td>for he was afraid of his sharp horns.</td>
<td>$P_5$ afraid(lion, horns(bull)) ba(S$<em>5$); ea(S$</em>{13}$)</td>
</tr>
<tr>
<td>$TE_5$</td>
<td>Hunger, however, presently compelled him to do something</td>
<td>$P_6$ compelled(lion, act) ba(S$_5$); ea(S$_5$)</td>
</tr>
<tr>
<td>$P_7$ hungry(lion) ba(S$<em>1$); ea(S$</em>{13}$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$TE_6$</td>
<td>and as the use of force did not promise success</td>
<td>$P_8$ believe(lion, ~promise(force, success)) ba(S$<em>4$); ea(S$</em>{13}$)</td>
</tr>
<tr>
<td>$TE_7$</td>
<td>he determined to resort to artifice</td>
<td>$P_9$ plan(lion, artifice) ba(S$<em>4$); ea(S$</em>{13}$)</td>
</tr>
<tr>
<td>$TE_8$</td>
<td>Going up to the Bull in friendly fashion</td>
<td>$P_{10}$ approach(lion, bull) ba(S$_5$); ea(S$_5$)</td>
</tr>
<tr>
<td>$P_{11}$ friendly(lion) ba(S$<em>5$); ea(S$</em>{13}$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$TE_9$</td>
<td>he said to him, “I cannot help saying how much I admire your magnificent figure.”</td>
<td>$P_{12}$ say(lion, bull, admire(lion, figure(bull))) ba(S$_5$); ea(S$_5$)</td>
</tr>
<tr>
<td>$P_{13}$ say(lion, bull, magnificent figuraure(bull))) ba(S$_5$), ea(S$_5$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$TE_{10}$</td>
<td>“What a fine head!”</td>
<td>$P_{14}$ say(lion, bull, fine(head(bull))) ba(S$_6$); ea(S$_7$)</td>
</tr>
<tr>
<td>$TE_{11}$</td>
<td>“What powerful shoulders and thighs!”</td>
<td>$P_{15}$ say(lion, bull, powerful(shoulders(bull)))) ba(S$_7$); ea(S$_8$)</td>
</tr>
<tr>
<td>$P_{16}$ say(lion, bull, powerful(shoulders(bull)))) ba(S$_8$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$TE_{12}$</td>
<td>“But, my dear friend, what in the world makes you wear those ugly horns?”</td>
<td>$P_{17}$ ask(lion, bull, reason(wear(bull, horns(bull)))) ba(S$<em>8$); ea(S$</em>{10}$)</td>
</tr>
<tr>
<td>$P_{18}$ say(lion, bull, ugly(horns(bull))) ba(S$<em>{10}$); ea(S$</em>{11}$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$TE_{13}$</td>
<td>“Believe me, you would do much better without them.”</td>
<td>$P_{19}$ say(lion, bull, ifThen(~have(bull, horns(bull))), succeed(bull))) ba(S$<em>{11}$); ea(S$</em>{12}$)</td>
</tr>
<tr>
<td>$TE_{14}$</td>
<td>The Bull was foolish enough to be persuaded by this flattery to have his horns cut off</td>
<td>$P_{20}$ foolish(bull) ba(S$_{12}$)</td>
</tr>
<tr>
<td>$P_{21}$ persuade(lion, bull, allow(bull, cutOff(helper, horns(bull)))) ba(S$<em>{12}$); ea(S$</em>{13}$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$TE_{15}$</td>
<td>and, having now lost his only means of defense,</td>
<td>$P_{22}$ lose(bull, ability(bull, defend(bull)))) ba(S$<em>{13}$); ea(S$</em>{15}$)</td>
</tr>
<tr>
<td>$TE_{16}$</td>
<td>fell an easy prey to the Lion.</td>
<td>$P_{23}$ identity(bull, prey(lion)) ba(S$<em>{14}$); ea(S$</em>{15}$)</td>
</tr>
</tbody>
</table>

Table 3.4: A textual- and timeline-layer encoding of “The Wily Lion”. 
telling time and the vertical axis is story time [Eco, 1995]. For each Proposition node on the timeline, we plot a point at \((x, y)\) where \(x\) is the ordinal position of the first Text node linking to the P node with \(ia\), and \(y\) is the time index of the State node linked by the P node with \(ba\). If the curve increases monotonically, each new span in the discourse advances the story time (a linear telling). A flat curve indicates that multiple discourse spans describe a single point in story time (a suspension of time). Certain other curves indicate that the narrator is “flashing back” or forward, or moving quickly or slowly through a period of story time (the so-called “tempo” of a story) [Mani, 2010].

Figure 3.8 gives four such plots, three for hypothetical stories and one for “The Wily Lion”. For every Text node along the horizontal axis, there are one or two bars for linked Proposition nodes. (There is usually one bar, but if a single span maps to two P nodes, the second is plotted in an adjacent grey bar.) Clockwise from the bottom left are: A “slow” story, in which the narrator describes several moments in detail; a “fast” story, where nearly every span invokes a new time state; a “flashback” story, where the narrator begins in the middle of the story and interrupts the flow of time to give a scene of background information; and on the bottom right, “The Wily Lion” as modeled in Table 3.4. In the latter case, temporal modeling alone (divorced from information about goals and plans) suggests thematic content. The first half of the fable moves slowly through time, detailing a moment with a set of statives relating to the lion. He is established as the protagonist; his mental states in a particular moment dominate the attention of the storyteller. In the second half of the fable, time moves swiftly until the story’s conclusion. The lion begins to act externally until the bull performs his only actions as agent in \(TE_{14}-TE_{15}\). The story is one of thinking, then doing.
Alternate Timelines

We have seen the “Reality timeline” used as a model context for textbase happenings in the timeline layer. Additional modalities are represented by separate Timeline and Proposition nodes. These *alternate timelines* indicate hypothetical and imagined modalities. Through the use of *referenced by* arcs, they take functional roles in Reality-timeline P nodes. For example, at a particular moment an agent may express speculation that some action will happen in the future. In our timeline encoding of “The Wily Lion”, node \( P_2 \) depicts a thought process by the lion about the royal feast the bull “would make” in a potential future. The feasting on the bull does, in the end, happen, but \( P_2 \) does not itself jump forward in story time because it concerns the lion’s mental state in the “present.” We used a predicate *potentialFuture* for this scenario. With an alternate timeline we can instead create a “scope of time” that represents speculation, and use an arc to refer to it in \( P_2 \). This feature adds formality to the model’s representation of imagined, desired or feared events, which are common in thematically rich stories.

Hypothetical events and states are particularly important to character agency, since goals and plans are themselves potential futures which may or may not come to pass. Hypothetical *past* states are also thematically important, as agents sometimes try to reason about the history of the story-world. In a mystery story, the detective has a goal for the future (to solve the case), but that goal is to develop hypotheses about *prior* events based on a collection of evidence.

To reiterate, the SIG allows additional Timeline nodes that have their own sets of State and Proposition nodes. We integrate these into the larger graph with two additional arcs:

**Referenced by** (r). Traverses from a Timeline node \( T \) to a Proposition node \( P \) that incorporates the timeline in a modal context. The timeline containing \( P \) is said to be a “parent” of the timeline represented by \( T \). An alternate timeline is, in a sense, a “inner scope” of narrative reality, existing wholly and exclusively in the context of the outer scope that references it. Timeline relationships are therefore tree-like, in that a timeline can have multiple children (inner scopes) but only one parent (outer scope). The Reality timeline is the root of the tree, and the “ancestor” of all alternate timelines:

\[
(P_1 \in T_1) \land r(T_2, P_1) \Rightarrow T_1 \neq T_2  \\
(P_1 \in T_1) \land r(T_2, P_1) \Rightarrow parent(T_1, T_2)  \\
parent(T_1, T_2) \lor (parent(T_1, T_3) \land ancestor(T_3, T_2)) \Rightarrow ancestor(T_1, T_2)  \\
ancestor(T_1, T_2) \Rightarrow \neg ancestor(T_2, T_1)
\]

**Equivalent** (e). If timeline \( T_1 \) is a parent of timeline \( T_2 \), a State node \( S_1 \in T_1 \) is equivalent to a State node \( S_1 \in T_2 \) if \( S_1 \) and \( S_1' \) are two modal contexts of the same functional State. That is, *equivalent* indicates that the same time slice is manifest as two State nodes in different timelines, and joins the nodes together as a common point of reference. The \( e \) arc can only join two states in different timelines that have an ancestral relationship. Multiple \( e \) arcs are permitted between the same two timelines, with the
Figure 3.9: Two configurations of alternate timelines in the timeline layer of a SIG.

<table>
<thead>
<tr>
<th>Node/Figure</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i)</td>
<td><strong>Henry thinks that Orson graduated from Cornell.</strong></td>
</tr>
<tr>
<td>State $S_1$</td>
<td>A moment of time.</td>
</tr>
<tr>
<td>Prop. $P_1$</td>
<td>Henry believes in the events of a separate scope of time.</td>
</tr>
<tr>
<td>Alternate Timeline</td>
<td>The scope of time believed by Henry in $P_1$.</td>
</tr>
<tr>
<td>State $S'_1$</td>
<td>Within Henry’s belief in $P_1$, the present moment of Henry’s belief.</td>
</tr>
<tr>
<td>State $S'_2$</td>
<td>Within Henry’s belief in $P_1$, a moment prior to $S'_2$ (Henry’s believing).</td>
</tr>
<tr>
<td>Prop. $P_2$</td>
<td>Orson graduates from Cornell.</td>
</tr>
<tr>
<td>(ii)</td>
<td><strong>Clarissa will have bought the flowers by 6 P.M.</strong></td>
</tr>
<tr>
<td>State $S_1$</td>
<td>A moment of time prior to 6 P.M.</td>
</tr>
<tr>
<td>State $S_2$</td>
<td>6 P.M.</td>
</tr>
<tr>
<td>Prop. $P_1$</td>
<td>The events in a separate scope of time occur.</td>
</tr>
<tr>
<td>Alternate Timeline</td>
<td>A scope of time.</td>
</tr>
<tr>
<td>State $S'_1$</td>
<td>Within the separate scope of time, the present moment equivalent to $S_1$.</td>
</tr>
<tr>
<td>State $S'_2$</td>
<td>Within the separate scope of time, a moment between $S'_1$ and $S'_3$.</td>
</tr>
<tr>
<td>State $S'_3$</td>
<td>Within the separate scope of time, a moment corresponding to 6 P.M.</td>
</tr>
<tr>
<td>Prop. $P_2$</td>
<td>Clarissa buys the flowers.</td>
</tr>
</tbody>
</table>

Logical constraint that the relative ordering must be preserved:

$$e(S'_1, S_1) \land e(S'_2, S_2) \land f(S_1, S_2) \Rightarrow f(S'_1, S'_2)$$

Figure 3.9 shows fragments of two example SIG encodings that invoke alternate timelines. The Reality timelines are depicted in white node boxes; the alternate timelines are drawn in grey. In 3.9(i), an action references a timeline in which a modal state $S'_2$ is equivalent to the action’s state $S_1$ in the Reality timeline. The modal state $S'_2$ is then preceded by another modal state $S'_1$ containing an imagined action $P_2$. Because the equivalent arc establishes a common point of reference between the two timelines, any action at $S'_1$ occurs “some time previous” to both $S'_2$ in the modal timeline and $S_1$ in Reality. This topology might represent that a character is thinking about an action that occurred at some point in the past (previous to the moment of thinking): “Henry thinks that Orson graduated from
Cornell” would be one example. 3.9(ii) similarly depicts a modal context for a possible future action $P_2$, because $P_2$ is attached to a modal state $S'_2$ which follows the common point of reference $e(S'_1, S_1)$. This figure, however, adds another temporal constraint by employing a second equivalent arc between $S'_2$, which follows $S'_2$, and $S_2$, which follows $S_1$. Because $P_2$ occurs between $S'_1$ and $S'_3$, in Reality it is imagined to occur between $S_1$ (the moment of imagining) and some future time $S_2$. A character may, in this case, be promising that some event happen will happen by a future deadline represented by $S_2$, e.g., “Clarissa will have bought the flowers by 6 P.M.”

Both of these examples use an $e$ arc to attach the modal timeline to the Reality timeline at a common point of reference. In the absence of any $e$ arcs, a modal timeline represents an entirely separate narrative scope with no points of attachment to Reality. This occurs when a story embeds a fictional inner story as told by a character (a “frame narrative”), such as One Thousand and One Nights and its serial storyteller Scheherazade. A nested story can be modeled as an embedded SIG encoding, with the storyteller-character taking the role of “focalizer” (narrating agent) [Bal, 1981; Bronzwaer, 1981; Genette, 1983]. The speech actions of Scheherazade become the discourse utterances of her own story; in essence, the timeline layer of the framing SIG encoding becomes the textual layer of the nested encoding.

Linguistically, alternate timelines allow us to model tenses and aspects in the discourse that refer to an ambiguous span of time in the Reality timeline. Consider the sentence: “John started to make breakfast but went to the store because he ran out of eggs.” It is semantically unclear whether John used his last egg during his breakfast preparation, perhaps dropping it, or if he used the last egg in some prior episode. A modal timeline such as Figure 3.9(i) preserves this ambiguity by asserting that an event (running out of eggs) takes place at some time in the past—the exact past time is unknown because there are no additional equivalence arcs to provide bounds. The past participle tense, “he had run out of eggs,” would allow us to draw such a bounding equivalence arc. It tells us that the “running out” event occurs prior to the “making breakfast” event.

Such a use of the equivalence arc is analogous to the notion of the reference time in Reichenbach’s [1947] study of tense and aspect. In Reichenbach’s approach, the temporal interpretation of a sentence is governed by the relative ordering of three important time points: the speech time S, the event time E, and a temporal point of reference R. This system maps onto the present model of alternate timelines. S is the point of attachment in the parent timeline (the state associated with the incoming referenced by arc), E is an event in the alternate timeline, and R is a time state in the parent timeline with an incoming equivalence arc. In 3.9(ii), S would be $S_1$, the speech time; E would be $P_2$, the event time, and R would be $S_2$, the reference time. This figure can be read in the future perfect tense given by the ordering S-E-R: “Clarissa will have bought the flowers” (by time $S_2$). This mapping assumes that a primary equivalence arc establishes a modal time state equivalent to the speech time, $S'_1$ in this case. In the absence of a secondary equivalence arc, R is set to be the same as S, resulting in a simple future tense with no separate reference time (“Clarissa will buy the flowers”). We will further investigate the relationship between alternate timelines and a model of tense and aspect in Section 4.4.
Discussion of the Representation of Time

We believe this approach to representing time is robust, in the sense that it is tolerant of partial encodings of the *fabula* timeline. Unlike temporal databases, we do not tie each state to a particular UTC timestamp or formally represent the relative lengths of time intervals signified by *followed by* (e.g., 9 hours 4 minutes passed between $S_1$ and $S_2$). In addition, if complete interval information cannot be inferred from the discourse, the timeline can be instantiated with *begins at* arcs alone (reducing the timeline to a set of points rather than intervals). The essential aspects of the SIG timeline are the relative orderings of TE and P nodes (via States and *begins at relations*).

This is not to say that reductionism is appropriate in all cases. Situations in which more precise information about time is relevant to the thematic content of the story are modeled in terms of that relevance. For example, the drama of the Puccini opera *Madama Butterfly* (with libretto by Luigi Illica and Giuseppe Giacosa) hinges on the long absence of Pinkerton, a U.S. Naval Officer, from Japan. Pinkerton has married a local girl, Cio-Cio San, who endures loneliness and financial hardship while waiting for Pinkerton’s ship to return. She gives up opportunities to remarry even though she does not hear from him for years. The drama derives not from the mere length of time that passes between the opera’s acts, as Cio-Cio-San waits, but from the decisions she takes to maximize her happiness based on a trust of Pinkerton’s intentions. When Pinkerton finally returns with an American wife in tow, the tragedy is not that an exact number of years has elapsed, but rather that a tremendous opportunity cost has been exacted from the heroine—Cio-Cio San’s plan has backfired and precluded her from finding happiness by any means. The large passage of time was the *enabler* of the affectual harm to Cio-Cio-San, but not the harm itself. That thematic idea, which is the purpose of the interpretative layer, can be encoded even though time is itself only represented as an ordering function.

In the next section, we transition away from the textual and timeline layers, and introduce the many aspects of the interpretative layer that allow us to model the tragedy of *Madama Butterfly* and other stories.

**3.3.2 Interpretative Layer**

The interpretative layer of the SIG depicts a situation model of the story-world. Like the timeline layer, it contains a set of nodes that represent story-world happenings. The major difference from the timeline layer is the manner in which these nodes are organized: Rather than by time, the interpretative layer takes a theory-of-mind (agent-centric) approach, structuring content by its motivational, intentional and affectual connections. We call the layer “interpretative” because the situation model is a subjective artifact that reflects a particular receiver’s interpretation of the story’s agents and their motivations. While we have developed annotation guidelines, the process of arriving at such an interpretation is not itself a part of the schemata.

Let us first define the final node types: Interpretative Proposition (I), Belief (B), Goal (G), and Affect (A). We call these the *interpretative nodes*.

**Interpretative Proposition (I).** The equivalent of P nodes in the interpretative context,
these nodes represent story-world content such as events and statives that may or may not have been expressed in the surface discourse.

**Belief (B).** A belief node acts as a frame, inside of which the content of other nodes is understood to be a state of the story-world in the mind of a discrete agent. This agent is an inherent and immutable attribute of the node (so that every Belief node that is instantiated is associated with an agent). This agent can be a single intentional entity or a set of entities who share the same beliefs. We use the notation $B:X()$ to describe a Belief frame, with $X$ referring to the agent in question, and the content of the frame appearing as a set of arguments. An unlimited number of interpretative nodes can be placed inside the frame. A belief frame can itself be negated to assert a lack of belief in its content (note the distinction between believing a statement $N$ is false, $B:X(\neg N)$, and not having the belief that the statement is true, $\neg B:X(N)$).

**Goal (G).** A goal node acts as a frame for other interpretative content, similar to a Belief. The difference is that the content of a Goal frame is understood to the state of the story-world as *desired* by the discrete agent. We notate Goals as $G:X()$.

**Affect (A).** An Affect node represents a baseline affectual state with respect to a discrete agent. As in Belief and Goal nodes, the agent can be a single intentional entity or a set of entities.

An in arc appears in the interpretative layer with a semantic meaning distinct from that of the in arc of the timeline layer:

**In (in), additional.** Traverses between an interpretative node and a Belief or Goal node representing the frame in which the interpretative node is situated. Each interpretative node must have 0 or 1 outgoing in arcs (each node can only belong to one frame), but a Goal or Belief node can have an unlimited number of incoming in arcs.

For the clarity of this discussion, we will draw Goal and Belief frames as graphical boxes that contain the nodes connected with in arcs (their content). Figure 3.10(i) depicts an example SIG encoding fragment that represents a two-part goal of agent X: for some action to happen, and for another agent Y to believe that some stative is true. Such a situation

---

**Figure 3.10: Nested agency frames, in two forms of graphical notation.**

(i)  
\[
\text{Goal: } X \\
\text{Action} \\
\text{Belief: } Y \\
\text{Stative}
\]

(ii)  
\[
\text{Goal: } X \\
\text{Action} \\
\text{Stative} \\
\text{Belief: } Y
\]
could be: “Larry wanted (to win the chess game against Debra, and have Debra believe (that he is a skilled player)).” 3.10(ii) shows the same graph fragment, but is drawn using the box notation. Note that “Action” and “Stative” nodes are, logically, both Interpretative Proposition (I) nodes; we label them more specifically for clarity.

Agency frames—goals and beliefs—can be nested indefinitely to model theory-of-mind interpretations of narrative meaning, as we saw in Section 3.2.2. This allows us to represent not only what agents believe about the world, but also what they believe each other’s beliefs, about each other’s beliefs about others’ beliefs, ad nauseam. When an interpretative proposition (I) node is not placed in any agency frame, it is in what we call the ground truth of the SIG: that which the narrating agent of the story asserts to be true in the scope of the story-world.4

The rest of this section is divided as follows: We first describe actualization, which allows the SIG to express changes in interpretative content over story time (such as when an outstanding goal is resolved through an outcome). We then introduce arcs to connect interpretative nodes into plans and attempts. Finally, we describe the manner in which Affect nodes may be used to express the affectual impact of interpretative content on certain agents.

### 3.3.2.1 Actualization

The interpretative layer, in and of itself, is timeless. As we noted in the introduction to this section, previous cognitive situation models conflate temporal and interpretative connections. The present model instead separates these two types of discourse relations, with temporal connections in the textual and timeline layers, and thematic content (goals, plans, and so on) in the interpretative layer.

However, time is still crucial to the interpretative layer. Goal outcomes, for example, must temporally follow attempts. The SIG assigns a temporal dimension to interpretative content by relying on its connections to the timeline layer to determine what interpretative nodes “happen,” and in what order. For each State node in the Reality timeline, a set of logical entailments determines which nodes in the interpretative layer are occurring at the corresponding point in story time, and which are not. This computation is called actualization.

For example, consider again the interpretative goal in Figure 3.10, in which Larry wants to win a chess game and have his opponent Debra believe that he is a skilled player. This figure has no sense of time—it is only a goal in isolation. But a timeline can actualize certain pieces of it in sequence, and draw a story out of it:

- At state $S_1$, the story is beginning.

No one has any goals. The entirety of Figure 3.10 is hypothetical (immaterial) at this point in time, because it does not yet exist in the story-world.

---

4The narrator may, of course, be unreliable [Booth, 1961]. Frame narratives are a particular risk, since the storyteller is itself a character who might be interested in distorting the facts. In a sense, all story-worlds are constructions of artificial realities, even when they purport to be non-fiction, because of the editing process inherent in the intentional act of storytelling [Genette, 1972]. For our purposes, all content placed in ground truth represents the objective reality of the story-world.
• At state $S_2$, Larry develops a goal to win a chess game against Debra and have her think he is a skilled player.

Larry goes from having no goal to having this particular goal. The goal frame becomes actualized. The goal content inside the frame (winning the game and having Debra believe he is skilled) is hypothetical, rather than actualized—it is a possible future.

• At state $S_3$, Larry wins the chess game against Debra.

The “winning” action inside the goal frame transitions from being hypothetical to being actual. The goal frame is still actual. Since Larry has a goal to win the chess game at $S_2$, and he does in fact win the chess game in $S_3$, he has achieved a positive outcome on this aspect of his goal. In general, this is the mechanism by which we express goal outcomes.

• Also at state $S_3$, Debra comes to believe that Larry is an unskilled chess player (perhaps she believes he has won the game unfairly).

The belief frame nested within Larry’s goal frame transitions from being hypothetical to being demonstrably false (what we call prevented/ceased). Debra’s opinion of Larry’s skills goes from undefined to “not skilled.” Larry has reached a negative (failed) outcome on this aspect of his goal.

This example demonstrates the logical property of the SIG whereby interpretative content has a certain actualization status relative to every state in the Reality timeline. The actualization status of the goal frame was hypothetical until it became “actualized” at $S_2$; the node representing the goal action itself (for Larry to win) was hypothetical until it was actualized at $S_3$; the frame representing the other aspect of the goal (Debra’s belief) was hypothetical until it was “prevented/ceased” in $S_3$.

In general, a node’s actualization status relative to some point in story time is always one of three conditions that describe the truth (within the story-world) of the node’s content at that time:

1. Hypothetical (H). The node’s content is in a hypothetical state, existing as a concept rather than as an assertion of a story-world happening. The present truth of a Hypothetical node is indeterminate; no assertion is made about whether the content is true within the story-world at the moment in question, or not.

2. Actualized (A). The node’s content is true (in effect; currently occurring in the story-world). Successful goal content is given A status at the point when it becomes successfully true.

3. Prevented/Ceased (PC). The node’s content is false (not in effect; decisively incompatible with the story-world). Nodes that are prevented/ceased not only are untrue at the present time, but given the current state of affairs, have been prevented from happening in the foreseeable future. In the language of prior models, a goal with a failed outcome has PC status.
Formally, we let the actualization status \( s \) of an interpretative node \( I \) at some time index \( n \) be one of \( \{H, A, PC\} \), and let every node’s status at time 0 be Hypothetical:

\[
\forall I \in I : \forall n \in N : s(I, n) \in \{H, A, PC\} \\
\forall I \in I : s(I, 0) := H
\]  

(3.7)  

(3.8)

Every interpretative node logically carries one actualization status for each state in the Reality timeline. In the example above, there are three states, \( S_1 \) to \( S_3 \), and four nodes of interpretative content. There are therefore 12 actualization statuses defined for this example, one for every node-state combination.

Actualization status transitions are **triggered** by particular arcs that traverse from the timeline layer to the interpretative layer. Using these arcs, the Reality timeline acts as an instruction set and clock for determining the actualization status of each node for each state. This is accomplished by virtue of the fact that Proposition nodes are totally ordered in story time, from the first P node in the earliest state (attached to the State node with no incoming \( f \) arcs) to the last node in the latest state. (For purposes of ordering propositions, only \( ba \) arcs are considered; \( ea \) arcs do not trigger changes in interpretative actualization.)

All interpretative nodes begin in Hypothetical (H) status. Then, each successive Proposition node associated with the Reality timeline has the opportunity to trigger a change in the actualization status of one or more interpretative nodes. There are two types of triggers: those that actualize \((\text{act})\) and those that prevent/cease \((\text{pc})\). These are not SIG arc types, but useful shorthands, as we will soon introduce multiple arc types that logically entail either \( \text{act} \) or \( \text{pc} \). To apply this to our chess example (see Figure 3.11(i)):
 CHAPTER 3. STORY INTENTION GRAPHS

<table>
<thead>
<tr>
<th>Prior Actualization Status</th>
<th>Incoming Trigger</th>
<th>New Actualization Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypothetical</td>
<td>Actualize</td>
<td>Actualized</td>
</tr>
<tr>
<td>Hypothetical</td>
<td>Prevent/ Cease</td>
<td>Prevented/ Ceased</td>
</tr>
<tr>
<td>Actualized</td>
<td>Actualize</td>
<td>Actualized</td>
</tr>
<tr>
<td>Actualized</td>
<td>Prevent/ Cease</td>
<td>Prevented/ Ceased</td>
</tr>
<tr>
<td>Prevented/ Ceased</td>
<td>Actualize</td>
<td>Actualized</td>
</tr>
<tr>
<td>Prevented/ Ceased</td>
<td>Prevent/ Cease</td>
<td>Prevented/ Ceased</td>
</tr>
</tbody>
</table>

Table 3.5: Transition of interpretative node actualization status upon receiving a trigger from a new time state.

- At state $S_1$, the story is beginning.
  There are no triggers.

- At state $S_2$, Larry develops a goal to win a chess game against Debra and have her think he is a skilled player.
  There is an arc that triggers *actualize* traversing from the motivating action at state $S_2$ to the node representing Larry’s goal frame.

- At state $S_3$, Larry wins the chess game against Debra.
  There is an arc that triggers *actualize* traversing from the “win” action at state $S_3$ to the “win” action inside the Larry’s goal frame.

- Also at state $S_3$, Debra comes to believe that Larry is an *unskilled* chess player.
  There is an arc that triggers *prevent/ cease* traversing from the “win” action at state $S_3$ to the “believe skilled” frame inside the Larry’s goal frame.

The actualization status of a node at a particular time state is a function of two factors: the node’s previous status (that is, its status with respect to the preceding P node), and the presence of any incoming trigger arcs. The effects of the two types of triggers are summarized in Table 3.5. *Actualize* triggers always cause the interpretative node to become Actualized, no matter the prior status; *prevent/ cease* triggers always cause the interpretative node to become Prevented/ Ceased. In other words, the truth-value of a node may alternate between Actualized and Prevented/ Ceased, or remain Hypothetical.

Formally, we let $a(P, I)$ and $pc(P, I)$ indicate that proposition node P triggers *actualize* and *prevent/ cease* (respectively) on some interpretative node I, such that both entail an actualization status for I with respect to the time index associated with P:

$$a(P, I) \land ba(P, S) \land t(S) = n \implies s(I, n) := A \quad (3.9)$$
$$pc(P, I) \land ba(P, S) \land t(S) = n \implies s(I, n) := PC \quad (3.10)$$

In the absence of any incoming arcs that entail such triggers, the actualization status of an interpretative node I at some time index $n$ is unchanged from the previous time index $n - 1$:

$$\forall n \in \mathbb{N} : \forall I \in \mathbf{I} : (\neg \exists P : ((a(P, I) \lor pc(P, I)) \land ba(P, S) \land t(S) = n) \implies s(I, n) := s(I, n - 1))$$
As we have seen, Figure 3.11 illustrates the triggering of actualization transitions by timeline P nodes. 3.11(i) shows a partial SIG encoding. The timeline layer includes two actions and the initial state $S_1$. The interpretative layer includes the same multi-part goal seen in Figure 3.10: Agent X (Larry) wants both for an action to happen (to win the game) and for Agent Y (Debra) to believe that some stative is true (that Larry is skilled). The actualization status of the goal at each state is shown in 3.11(ii-iv) by means of a suffix associated with each node: /H, /A, or /PC. Actualization statuses are also drawn graphically, with light shading for Actualized status (/A), dark shading for Prevented/Ceased status (/PC) and no shading for Hypothetical status (/H). The two timeline actions trigger three actualization status changes: there are two actualize triggers and one prevent/cease trigger. As these are not themselves SIG arcs, they are drawn with dashed arrows.

With respect to nesting, actualizations must proceed from the “outside in.” No node can be actualized or ceased if it is in a frame that still has Hypothetical status.

### 3.3.2.2 Actualizing Arcs

Let us now introduce the first four of the 13 SIG arc types which relate to the interpretative layer. These relations always connect timeline-layer nodes to interpretative-layer nodes, and are the only arc types that trigger actualization status changes. The first three of these are actualizing triggers; the last (c) is a preventing/ceasing trigger.

**Interpreted as (ia), additional.** Traverses between a timeline P node and an interpretative frame or I node when there is a direct equivalence (that is, the content of the interpretative node is a paraphrase of the content of the timeline node); the same arc is similarly used to connect equivalent nodes between the textual and timeline layers.

**Implies (im).** Traverses between a timeline P node and an interpretative frame or I node when the interpretative content can be inferred from the timeline content, but there is not a direct equivalence. This “weaker” form of ia connects a timeline happening with a node of interpretative content that it entails without stating outright. It should not be used if ia is possible.

**Actualizes (a).** Traverses between a timeline P node and any interpretative node (frame, I node or Affect node) when the interpretative content is actualized as a causal result. This “weaker” form of im and ia connects a timeline P node with a node of interpretative content that it causes, but does not either state or entail directly. It should not be used if ia or im are possible, but rather, when a timeline happening can be inferred to indirectly trigger an actualization.

**Ceases (c).** Traverses between a timeline P node and any interpretative node (frame, I node or Affect node) when the interpretative content is prevented/ceased as a causal result. Like a, it signifies that a timeline event implies a consequence in the interpretative layer.

Through I nodes and ia, im and a arcs, the model supports the representation of any fact that might be implied or stated by a timeline P node. This does imply that a proper
encoding is one that takes every opportunity to encode a consequence of an event, whether stated or unstated in the original discourse. That is, we wish to avoid the “frame problem” in artificial intelligence, in which an event has a prohibitively large number of possible consequences throughout the story-world to formally model [McCarthy and Hayes, 1969]. In a SIG, the task is not to model interpretative propositions for every consequence of an event, but for only those consequences that significantly impact the thematic content of the story. The test for “impact significance” has to do with the Affect nodes we will soon introduce: Any consequence that does not impact the affectual state of at least one agent in a manner that the storyteller explicated or implied in the story should not be encoded as a node.

Figure 3.12 shows a possible interpretative encoding for a small section of the “Wily Lion” timeline, as outlined in Table 3.4. The action at $S_1$, in which the lion watches the bull feed from the meadow, actualizes two nodes: the frame indicating that the bull wants to feed from the meadow, and the interpretative action that indicates his successful feeding. The initial story state at $S_1$ is one in which there is one goal, and it is already being successfully fulfilled. The action at $S_2$ implies that the lion has conceived of a goal to eat the bull; $S_{14}$ triggers both an actualization of the lion’s goal content (a successful outcome for the lion) and a cessation of the bull’s goal content (a loss for the bull). The overall dramatic arc in this encoding of the story is one of a tradeoff—the lion’s goal is satisfied by the same action that ceases the satisfaction of the bull’s goal.

3.3.2.3 Plans

We have seen actualization triggers indicate goal outcomes. However, we earlier saw that such outcomes are only part of a thematic narrative. Our schemata also needs a representation for a strategy toward fulfilling a goal—a plan, with subgoals that make progress toward actualizing the larger goal. In the case of “The Wily Lion”, the plan is a multi-stage scheme on the part of the lion to take advantage of the bull’s vanity, so that the bull takes action which removes (ceases) the horns which serve as an obstacle in the way of a hot lunch. The plan is never stated in the text; we are told that the lion decides to use artifice, but the
details of the intended artifice are never made explicit. It is up to the reader to infer what the plan is. Our schemata allows a receiver to encode not only his or her inference of the lion’s plan, but the gradual reveal of that plan by the narrator as the discourse unfolds. The next set of relations provide a mechanism for describing the strategies and possible futures of each agent, whether implied or explicit.

A plan is modeled in the interpretative layer as a chain of connected nodes inside a Goal frame. Each node is a “subgoal” that leads to the ultimate goal at the end of the chain. The chain is connected with directed arcs that indicate expected causality: the agent believes that the actualization of one subgoal would lead to the actualization of the superordinate goal that lies next on the chain, and so on, leading to the ultimate goal. Crucially, these expectations are themselves beliefs of the agent. These beliefs may be mistaken. For instance, an agent may devise a plan to bring about rain by praying to rain gods, even though in the ground truth of the story-world, no causal connection exists between the acts of praying and raining.

As we mentioned, each interpretative node begins with Hypothetical (H) status. This is true for each of the subgoal steps of a plan as well. Just as a single goal is understood to have a successful outcome when it is actualized, a plan is a multi-stage goal where each step can be individually actualized when (and if) it is achieved. Similarly, a plan that fails can be ceased at the point of failure.

The relations that define expected and/or intended futures are:

**Would Cause (wc).** Traverses from one interpretative node to another interpretative node. Signifies that in the belief context of the originating node, an actualization of the originating node would causally lead to (is sufficient for) an actualization of the destination node.

**Would Prevent (wp).** Traverses from one interpretative node to another interpretative node. Signifies that in the belief context of the originating node, an actualization of the originating node would causally lead to (is sufficient for) a prevention/cessation of the destination node.

**Precondition for (pf).** Only differs from *would cause* in that it signifies that an actualization of the originating node is necessary for an actualization of the destination node (but not necessarily sufficient).

**Precondition against (pa).** Only differs from *would prevent* in that it signifies that an actualization of the originating node is necessary for a prevention/cessation of the destination node (but not necessarily sufficient).

As an illustrative example, Figure 3.13 shows an interpretation of the lion’s plan. The actualization statuses are drawn with respect to state $S_5$ in the timeline laid out in Table 3.4, when the lion approaches the bull. At this point in the story, the lion has actualized a mental state in which he lays out a plan consisting of a seven-step causal chain. The lion plans to instill a goal on the part of the bull, which (the lion believes) would cause the bull to have his horns removed. The removal of the horns would allow the lion to kill and eat the bull.
As we mentioned earlier, applying theory of mind to literature suggests that much planning involves the management of the goals, plans and beliefs of others. In this example, a nested goal frame acts as a step in a plan, a subgoal that must be actualized in the same manner as if it were an I node. The lion’s plan calls for the bull to construct his own plan in which removing his horns is the first step. When the bull takes this action, it implies that the bull has indeed decided to embark on such a strategy (i.e., he has actualized the inner goal frame), and the lion can proceed with the next step of his larger plan.

Note in particular the duplication of remove(bull, horns) in two contexts. It exists in the bull’s plan as a means to making the bull handsome in the lion’s view, but in the lion’s plan as a means for killing the bull. When the bull does have his horns removed, it actualizes both nodes and furthers both plans. As such, this is a graph topology that depicts an ulterior motive or a hidden agenda.

In general, it is not necessary for all elements of a plan to be within the same structural frame (that is, all connecting to the same Goal node with in arcs). The meaning would be the same if the last four elements of the chain (from dangerous(bull) to eat(lion, bull)) were in a separate goal frame of the lion’s. In this case, the would prevent arc traversing to dangerous(bull) would cross from one goal frame to another. Plan chains may also involve segments in ground truth; only the beginnings and ends of plans must be inside goal frames.

In general, a plan can include not only the sequential actualization of I nodes, but the deliberate cessation of a node which is blocking the route to a goal via would prevent. Triggering prevent/cease on an actualized node that, through would prevent, blocks a desired state is a form of “double negation” that is equivalent to actualizing a node that would in turn actualize a desired state. Briefly put, a plan step can either be about ceasing an undesired state that is actualized, or actualizing a desired state that is hypothetical or ceased. In this particular case, the lion’s problem is that the bull holds a certain attribute, that it is well defended by its horns, and the actualized nature of that fact is preventing the
l lion from being able to kill and eat the bull. Only by ceasing the attribute, thereby cutting off the triggering of prevent/cease on his able() stative, does the lion gain the power to pursue his ultimate goal.

Would cause and would prevent arcs carry no logical constraints regarding the actualization statuses of either their source or their destination nodes at any state in the timeline. They do, however, imply that the agent expects the actualization status of the destination node to change once the actualization status of the source node changes. This expectation may be violated, and that violation may be a crucial dramatic turning point. Aristotle [1961] defined this as peripeteia, the point in a tragedy when the hero suffers a reversal of fortune after his expectations are violated. Peripeteia often goes hand-in-hand with anagnorisis, when the hero undergoes a revelation about himself and his situation. In Figure 3.13, the lion’s plan is predicated on the expectation that actualizing remove(bull, horns) would cease dangerous(bull). The lion may have found that, contrary to his expectation, the bull continued to be a formidable opponent without horns—and therefore he was the one who had been tricked while attempting to be the trickster. Such an outcome would be an example of both anagnorisis and peripeteia. We further explore the capability of the SIG to represent these concepts in Appendix B.

The distinctions between would cause/prevent and precondition for/against, respectively, are that satisfying a precondition does not cause the agent to expect the actualization status of the destination node to change—the agent believes that the preconditions are necessary but not necessarily sufficient. We use precondition for for the last two of the lion’s plan steps because they are about enablement; whether he chooses to exploit this ability once it is actualized is up to him. In general, the precondition arcs are useful for when a goal requires multiple parallel plans. In “Little Red Riding Hood,” the wolf seeks to fool the girl into believing that he is her grandmother by succeeding in two parallel tasks: disguising himself as the grandmother in appearance, and feigning the grandmother’s voice. Both are preconditions but neither is sufficient for the girl to lower her guard.

As we have mentioned, a plan may include steps which are inferred by the reader, in that there are no equivalent textual-layer or timeline-layer nodes. In this case, it is never stated in the original story that the lion’s plan is to trigger a plan on the part of the bull. This inference is enabled by world-knowledge and mind-reading processes that are not themselves a part of the descriptive SIG schemata.

It is technically possible that the lion has an altogether different plan at $S_5$ than the one depicted in Figure 3.13. The illustrated plan assumes that all the events following $S_5$ transpire more or less as envisioned by the lion at $S_5$. It is possible, however unlikely to us, that the lion sincerely wishes for the bull to become handsome because that would satisfy his hunger by other means (such as by increasing tourism and economic activity in their corner of the plains), and that when the bull removes his horns, the lion unfortunately succumbs to the baser instincts he has been repressing in a bid to remain acceptable to the civilized world. Such a reading is more about the lion’s internal conflict with his moral compass than it is about the bull’s foolishness. In a larger sense, stories rarely explicate the mental states of all their agents at all times; most, like this one, explicate some mental states and strongly imply others through action. Some narratives deliberately leave intentions and beliefs ambiguous. Our approach allows an annotator to encode his or her own reading of
the entire story, including both explicit and implicit thematic content. Both types of content can then become data for automatic processing such as the identification of analogies. (We attempt this in Chapter 5. While multiple encodings can also represent plural readings by the same annotator, each with alternative inferences, we asked annotators to settle on a single, preferred reading for this experiment.)

3.3.2.4 The Inference of Causality

The SIG can represent actualized causality between two I nodes A and B (and, by extension, the timeline P nodes that actualize or prevent/cease them) in four ways. These follow intuitively from the definitions of precondition for/against as “necessity” relationships and would cause/prevent as “sufficiency” relationships:

1. A is newly actualized, B is newly actualized at the same or a following time state, and A would cause B (Figure 3.14(i)). A caused B, in whole or in part.

Example: “Going to the loud concert would give Thomas tinnitus. Thomas went to the loud concert, so he got tinnitus.”

2. A is newly actualized, B is newly prevented/ceased at the same or a following time state, and A would prevent B (Figure 3.14(ii)). A caused the prevention/cessation of B, in whole or in part.

Example: “Going to the concert would prevent Adam from getting to work on time the next day. Adam went to the concert, so he failed to go to work on time the next day.”
CHAPTER 3. STORY INTENTION GRAPHS

3. A is newly prevented/ceased, B is newly actualized at the same or a following time state, and A precondition against B (Figure 3.14(iii)). The prevention/cessation of A allowed B to happen, in whole or in part.

Example: “Nathan’s excellent social skills, among his other abilities, kept him from losing his job. When he became extremely antisocial, his company decided to let him go.”

4. A is newly prevented/ceased, B is newly prevented/ceased at the same or a following time state, and A precondition for B (Figure 3.14(iv)). The prevention/cessation of A allowed the prevention/cessation of B to happen, in whole or in part.

Example: “The financier’s support was an integral part of the art gallery’s operational budget. The financier pulled his support, so the art gallery folded.”

See Appendix C.2 for formal descriptions of these four scenarios. Note that they are symmetric: We may take 3.14(i), invert B and its incoming arcs, and arrive at 3.14(ii) as an alternate formulation of the same underlying relationship (e.g., “Going to the loud concert would end Thomas’s run of not getting tinnitus.”). Also note that for illustrative purposes, Figure 3.14 varies the temporal relationship between A and B. In 3.14(i) and 3.14(ii), A and B are linked to separate timeline P nodes in sequential time states. In 3.14(iii), A and B are linked to separate nodes in the same time state. In 3.14(iv), A and B are linked to the same P node. All three temporal scenarios allow the inference of causality from A to B. However, the flow of time cannot be reversed; no causality can be inferred if, in any of the five scenarios, B begins at a state preceding that of A’s onset state.

These examples demonstrate the representation of “ground truth” with respect to the causal relationships between events. In addition, as we mentioned above, the causality arcs can be used inside certain belief contexts (that is, agency frames). In this case, they represent an agent’s belief about the causal relationship between two hypothetical or actual events—a belief which may or may not be true with respect to the ground truth. For example, an agent may be mistaken or ignorant about what would happen if it tries to execute a plan, or it may draw a false conclusion about the causal antecedent of an event. We give example encodings of such scenarios in Appendix B.

3.3.2.5 Epistemology of Belief Frames

In the interpretative layer, both nodes and frames can have incoming arcs that trigger actualize or prevent/cease. The actualization of a frame refers to the mental state of the agent associated with the frame, but not necessarily to the content found within the frame. For instance, the actualization of a belief frame that proposition A is true says nothing about whether A is indeed true—it only asserts that the agent believes A. In the chess example from Figure 3.11, we triggered prevent/cease on Debra’s belief that Larry is a skilled player. This denotes that Debra does not believe Larry is skilled, but makes no logical assertion about whether Larry is skilled or unskilled. The I node regarding Larry’s attribute as a skilled player is left with Hypothetical (indeterminate) status at the end of the timeline.

There are three ways to logically indicate that the agent is correct or incorrect in its belief in an assertion A:
<table>
<thead>
<tr>
<th>Diagram</th>
<th>Example</th>
<th>( E_1 )</th>
<th>( E_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i)</td>
<td>Andy thought that April was in Toronto.</td>
<td>Stative ( E_1 )</td>
<td>Done by agent ( X ), not materialized</td>
</tr>
<tr>
<td>(ii)</td>
<td>Andy thought that if April was in Toronto, she could not come to his birthday party.</td>
<td>Stative ( E_1 )</td>
<td>Done by agent ( X ), would cause ( E_2 )</td>
</tr>
<tr>
<td>(iii)</td>
<td>Andy thought that April was in Toronto, and therefore could not come to his birthday party.</td>
<td>Stative ( E_1 )</td>
<td>Done by agent ( X ), would cause ( E_2 )</td>
</tr>
<tr>
<td></td>
<td>Caroline thought it would rain within the hour.</td>
<td>Done by agent ( X ), not materialized</td>
<td>Materialized</td>
</tr>
<tr>
<td></td>
<td>Caroline thought that if it rained within the hour, she would have to reschedule her arboretum tour.</td>
<td>Done by agent ( X ), would cause ( E_2 )</td>
<td>Materialized</td>
</tr>
<tr>
<td></td>
<td>Caroline thought it would rain within the hour, in which case she would have to reschedule her arboretum tour.</td>
<td>Done by agent ( X ), would cause ( E_2 )</td>
<td>Materialized</td>
</tr>
</tbody>
</table>

Figure 3.15: Belief frames in a SIG can refer to (i) an agent’s belief in a proposition such as a stative, (ii) an agent’s belief in the hypothetical relationship between two propositions, or (iii) the combination of (i) and (ii) with respect to a single proposition.

1. Actualize or prevent/cease (respectively) the node containing \( A \) itself, inside the belief frame;
2. Actualize or prevent/cease (respectively) an interpretative node with identical content \((A)\) in ground truth; or
3. Prevent/cease or actualize (respectively) an interpretative node with negated content \((\neg A)\) in ground truth.

Once actualized, the semantic meaning of the belief frame depends on the structure of the subgraph found within the frame. If an I node inside a belief frame has an outgoing arc, the representation means that the agent believes the relation rather than the assertion in the node, and the actualization of the frame makes no commitment to whether the assertion is true or whether the agent believes the assertion to be true. Figure 3.15 illustrates an I node in a belief frame with and without an outgoing arc. In Figure 3.15(i), the stative \( E_1 \) is believed by Agent \( X \); in Figure 3.15(ii), agent \( X \) believes that there is a causal relationship between hypothetical statives \( E_1 \) and \( E_2 \) but is not depicted to believe that either is true or false; finally, in Figure 3.15(iii), the agent believes both that \( E_1 \) is true and that \( E_1 \) would cause \( E_2 \). Figure 3.15(iii) depicts what we call an expectation of \( X \) that \( E_2 \) is actualized by a subsequent timeline happening. None of these three examples commit to whether \( E_1 \) is ever true, whether \( E_2 \) is ever true, or whether \( E_1 \) would truly cause \( E_2 \).
3.3.2.6 Attempts

We have seen four types of SIG arcs that connect the timeline and interpretative layers (interpreted as, implies, actualizes and ceases). There are two others. These deal with the intentionality of an agent with respect to a goal or a plan:

**Attempt to cause (ac).** Traverses between a timeline P node with an agent and an interpretative frame or node when it is to be understood that the timeline happening is performed by its agent as an intentional attempt to bring about the actualization of the interpretative content, such that any actualizations/cessations by the timeline node are unintended by the agent unless they are also connected to the node with an attempt arc.

**Attempt to prevent (ap).** Traverses between a timeline P node with an agent and an interpretative frame or node when it is to be understood that the timeline happening is performed by its agent as an intentional attempt to bring about the prevention/cessation of the interpretative node, such that any actualizations/cessations of the timeline node are unintended by the agent unless they are also connected to the node with an attempt arc.

Neither of these arcs triggers a change in the actualization status of the interpretative node to which it connects. They do, however, signal that the agent in a timeline node is knowingly and willfully acting to try to bring about such a change in actualization status. These arcs are analogous to the Attempt (A) nodes in Trabasso's cognitive model; in a larger sense, they are motivated by the experiments we saw in Section 3.2.1 that actions executed in pursuit of goals are a key part of memory retention for goals themselves and for stories overall.

In much of the present fable, many of the actions are understood as being attempts to actualize a part of a plan. The timeline nodes $P_{10}$ through $P_{19}$ in Table 3.4 show the lion approaching the bull and speaking to him. With respect to the lion’s plan in Figure 3.13, these actions do not themselves actualize or cease any interpretative nodes. They are, however, related to the plan, in that they are all an attempt to cause the hypothetical first step (flatter(lion, bull)).

3.3.2.7 Affect

The final aspect of the interpretative layer, and our overall schemata, is the Affect node. This node is designed to represent an ultimate answer to the question of why an agent acts.

In a representation without Affect nodes, “why” can be answered with plans that generalize to larger and larger purposes. We say that the lion tricks the bull in order to be able to kill it, so that it can eat it; unfortunately, this does not intrinsically represent the meaning of eat() to either the predator or the prey. We are concerned with how an action fits into an overall plan, but in the case of eat(), the “plan” is simply commonsense biology (the lion wants to eat in order to digest the bull, which gives the lion a source of protein, which is metabolized into energy, which is expended to sustain basic life functions, and so on). None of these nominally superordinate goals are necessary for understanding the
### Table 3.6: Interactions between actualization status transitions and arcs relating to Affect nodes.

<table>
<thead>
<tr>
<th>Prior Status</th>
<th>New Status</th>
<th>Arc</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypothetical</td>
<td>Actualized</td>
<td>Provides For</td>
<td>The agent achieves something positive</td>
</tr>
<tr>
<td>Hypothetical</td>
<td>Prevented/Ceased</td>
<td>Provides For</td>
<td>The agent is hit by something negative</td>
</tr>
<tr>
<td>Actualized</td>
<td>Prevented/Ceased</td>
<td>Provides For</td>
<td>The agent loses something positive</td>
</tr>
<tr>
<td>Prevented/Ceased</td>
<td>Actualized</td>
<td>Provides For</td>
<td>The agent is freed of something negative</td>
</tr>
<tr>
<td>Hypothetical</td>
<td>Actualized</td>
<td>Damages</td>
<td>The agent is hit by something negative</td>
</tr>
<tr>
<td>Hypothetical</td>
<td>Prevented/Ceased</td>
<td>Damages</td>
<td>The agent avoids something negative</td>
</tr>
<tr>
<td>Actualized</td>
<td>Prevented/Ceased</td>
<td>Damages</td>
<td>The agent is freed of something negative</td>
</tr>
<tr>
<td>Prevented/Ceased</td>
<td>Actualized</td>
<td>Damages</td>
<td>The agent is hit by something negative</td>
</tr>
</tbody>
</table>

story. At some point, a goal’s rationale must be general enough that it is recognizable as a constant aspect of the human condition. Eating is necessary for health, and health “just is” as a rationale for action. We simply wish to encode that the lion’s eating of the bull is ultimately good for the lion, in terms of affectual impact, and bad for the bull.

Affect nodes fulfill this purpose by acting as affectively-charged termini for plans. They represent our premise that thematic content is ultimately about the relationships between agents and their basic needs. For each interpretative node, they answer the question: Why is this node ultimately relevant to an agent? Why is this aspect of the story interesting as an aspect of a tellable narrative? In short, Affect nodes are the way we represent the affectual impact of each interpretative node with respect to each agent.

As a SIG node, Affect (A) can be instantiated in the interpretative layer an unlimited number of times, but each instance must be connected to the larger graph structure. Each instance includes two features, an agent and a type; we will describe types in a moment. Affect nodes are connected to interpretative nodes and frames by one of the two following arcs:

#### Provides for (p).
Traverses from an interpretative node to an Affect node. Signifies that in the belief context of the interpretative node, the actualization of that node implies a positive affectual impact on the Affect node’s agent in a manner consistent with its type, and the prevention/cessation of the interpretative node has a corresponding deleterious impact.

#### Damages (d).
Traverses from an interpretative node to an Affect node. Signifies that in the belief context of the interpretative node, the actualization of that node implies a deleterious affectual impact on the Affect node’s agent in a manner consistent with its type, and the actualization of the interpretative node has a corresponding positive impact.

Semantically, Affect nodes represent the basic needs of agents as conscious entities. To provide for an Affect node is to positively affect the agent in question; to damage an Affect node is to negatively affect the agent. Table 3.6 considers an interpretative node that relates to an Affect node and lists the effects of changing the actualization status of the
interpretative node. In essence, actualizing a node which provides for an Affect node helps
the agent in question, and preventing/ceasing the node hurts the agent in question.

Logically, an encoding can be read “backward” from Affect nodes to understand the
affective context of nodes that are not directly connected. For instance, in Figure 3.16(i),
Agent X has a goal in which E would cause F, which itself provides for an Affect node. E
thus has an indirect but positive affectual impact on X. (Our drawings depict Affect nodes
as non-frame labels with white text over a black fill.) When X attempts to cause E, X
is also attempting to cause F and, in turn, actualize the Affect node. B, E and F are all
oriented as being ultimately about a positive affect for X. We call this type of inference
goal closure. In Appendix C.1, we give a set of formal rules for inferring the affective
meaning of a node based on an arbitrary number of “hops” to an Affect node.

Not everything that transpires in a story deserves to be oriented toward a particular
agent’s affectual state. A description of a space, for instance, might serve no other purpose
than to set the scene for the reader. Accordingly, our schemata does not call for Affect
nodes to be attached to every interpretative node. To illustrate this, Figure 3.16(i) includes
two interpretative propositions that exist in ground truth: C, which has no affectual impact,
and D, which has a deleterious affect on an agent by linking to an Affect node with damages.

However, in a proper SIG all content inside goal frames must relate to one or more
Affect nodes, either directly or through closure. Closure is possible for a node if there is
a path in the graph leading from the node to an Affect node. The path must only follow

---

Figure 3.16: Legal SIG encoding (top) and one that violates Affect node usage.
some combination of \textit{in, would cause, would prevent, precondition for, precondition against, provides for} and \textit{damages}. This rule has a graphical interpretation which stipulates that all content inside a goal frame must \textbf{drain} to an Affect node by following arcs of these seven types until it reaches one—that is, for each goal node, one should be able to trace a path from the timeline through that goal node to an Affect node while only following forward arcs. A node inside a goal frame for which there is no path to an Affect node violates the schemata. The purpose of this rule is to ensure that every goal is annotated with the affectual impact that motivates it. Figure 3.16(ii) illustrates an illegal encoding: Interpretative Proposition L, in the goal frame for Agent Y, does not have any outgoing arcs through which it could drain, and so no possible affectual impact can be ascribed to it.

Another restriction on goal closure is that \textbf{no cycles} are allowed. That is, the subgraph of a SIG encoding that only includes these seven arc types and the nodes that are incident to them must be a directed acyclic graph (but not necessarily a connected one). This is because a goal closure arc implies both temporal and causal ordering. The causal path of a hypothetical plan must always propagate forward to an Affect node. 3.16(ii) is also an illegal encoding because nodes K and J \textit{would prevent} each other, forming a causal cycle. One might suggest such an encoding in order to convey the concept of two mutually exclusive goals, such as in a conflict between two agents. The proper encoding for this scenario, demonstrated in Section B.5, uses two additional plan steps (one for each frame) to convey the same relationship without causing a cycle.

Affect nodes always represent the “ground truth” of affectual impact on an agent. They cannot be placed inside goal or belief frames with \textit{in} (though for graphical convenience, we sometimes draw them inside frame boxes). Strictly speaking, Affect nodes are not the endpoints of agent-intended plans, but metadata provided about each plan. This does not limit the expressibility of the schemata in terms of agent beliefs about affect. One may still encode a scenario where an agent expects an event to cause a positive affectual impact, only to have it cause a negative impact (Figure B.5).

### 3.3.2.8 Affect Typing

We mentioned earlier that P nodes in the timeline and I nodes in the interpretative layer can be cross-indexed with other annotation schemes applied to the same discourse. For example, a propositional encoding of a span of text, with semantic role labeling, can be associated with a Proposition node (hence its name). Only the identity of the agent (if any) is strictly necessary metadata for each node. The same is true of Affect nodes. Each instantiation of an Affect node can be cross-indexed with a knowledge representation for the “type” of affectual impact represented by the node. This can be useful when multiple Affect nodes are used for different purposes according to the semantics of the narrative. For instance, one may use two Affect nodes in a scenario where a single event is good in one manner and bad in another manner (a trade-off—see Figure B.2).

We make no claim in this thesis as to what knowledge representation is best for Affect nodes with respect to any narrative corpus, including Aesop’s. We only claim that such typing can increase the expressive range of the schemata. However, for purposes of demonstration, we have devised and implemented a set of types based on prior investigations into
the psychology of human motivation. The most well-known set of types is a hierarchy of needs devised by Maslow [1943]. To Maslow, “practically all organismic states are to be understood as motivated and as motivating,” a sentiment which dovetails with our notion of goal closure. He identifies five broad categories: physiological needs (those needed to maintain bodily homeostasis, such as food and sleep), safety needs (protection from wild animals, extremes of temperature, criminals, etc.), and the needs for love (affection and belongingness), esteem (self-confidence and the respect of others), and self-actualization (the fulfillment of one’s potential). These categories are hierarchical in that one tends to not be a concern unless the previous ones are satisfied. In another classification, Max-Neef [1992] devises an ontology of needs along two interacting dimensions: existential (being, having, doing and interacting) and axiological (subsistence, protection, affection, understanding, participation, creation, leisure, identity and freedom). Max-Neef argues that these basic needs are not only few, finite, and classifiable, but consistent across cultures and through historical periods. We have adapted these typings into the following twelve Affect types found in Table 3.7.

Although these categories are distinct, they are non-exclusive. One action may simultaneously cover multiple types. Such an action would be connected to multiple Affect nodes. For example, a father caring for a sick son is acting both for his son’s health and for his own love. We presented these types to annotators in the experiments we will describe in Chapter 5. We notate Affect nodes with a period between the agent and the type (e.g., X.FREEDOM).

Figure 3.17 augments Figure 3.13, the plan diagram for “The Wily Lion”, by adding typed Affect nodes. The plan now represents the notion that the lion’s overall purpose is to help himself by providing for LION.HEALTH. His plan for doing this involves instilling a goal on the bull’s part to act toward his own positive ends. More specifically, the lion prompts the bull to act in such a way that would favor the bull’s ego. The same action,
<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
<th>Motivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Life</td>
<td>Continuation of basic life functions; existence vs. non-existence.</td>
<td>Subsistence</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[Max-Neef, 1992]</td>
</tr>
<tr>
<td>Health</td>
<td>Freedom from pain, disease, malnutrition, and other physical/mental ailments. (If a loss permits the character to live in greater pain, it is a Health matter; if life and death are immediately at stake, it is a Life matter.)</td>
<td>Safety</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[Maslow, 1943], Protection [Max-Neef, 1992]</td>
</tr>
<tr>
<td>Ego</td>
<td>A positive perception of one’s qualities by one’s self and by others.</td>
<td>Esteem</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[Maslow, 1943]</td>
</tr>
<tr>
<td>Wealth</td>
<td>Material possessions or currency, above that needed for basic sustenance (those for Health).</td>
<td>Esteem + Leisure</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[Max-Neef, 1992]</td>
</tr>
<tr>
<td>Love</td>
<td>Feelings of fondness, warmth, and romance for and from another person; familial companionship; compassion or a desire to heal the world.</td>
<td>Affection</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[Max-Neef, 1992]</td>
</tr>
<tr>
<td>Leisure</td>
<td>Entertainment and enjoyment, whether from peaceful solitude, active socializing, or another form of recreation.</td>
<td>Leisure</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[Max-Neef, 1992]</td>
</tr>
<tr>
<td>Membership</td>
<td>Feeling of belonging to a group; acceptance by its other members and holding one’s self positively by the norms and customs of that group. The group can be ethnic, social, economic, or in the micro sense, about cliques and clubs.</td>
<td>Identity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[Max-Neef, 1992]</td>
</tr>
<tr>
<td>Actualization</td>
<td>Fulfillment of one’s artistic, athletic, spiritual, professional or other aspirational potential in an elective endeavor.</td>
<td>Self-actualization</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[Maslow, 1943], Creation [Max-Neef, 1992]</td>
</tr>
<tr>
<td>Freedom</td>
<td>The state of being unrestricted in movement, action and behavior, whether the restricting force is other characters, natural forces, or an internal struggle.</td>
<td>Freedom</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[Max-Neef, 1992]</td>
</tr>
<tr>
<td>Justice</td>
<td>The perception that one’s code of ethics is being executed fairly; the desire to see good outcomes come to those whose actions one believes are moral.</td>
<td>Safety</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[Maslow, 1943]</td>
</tr>
<tr>
<td>Enlightenment</td>
<td>A more full and accurate view of the world, whether through education, spirituality or other means.</td>
<td>Understanding</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[Max-Neef, 1992]</td>
</tr>
<tr>
<td>Honor</td>
<td>The perception that one is fulfilling one’s own code of ethics, and that of the law and moral code of the community to which one belongs.</td>
<td>Participation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[Max-Neef, 1992]</td>
</tr>
</tbody>
</table>

Table 3.7: Affect typing used for the present study.
remove(bull, horns), advances toward three affect states when it is actualized: The bull believes it is helping the bull’s ego, but the lion knows it is enabling him to eat the bull—an action that the lion seeks for purposes of his health, but that has the side effect of ending the bull’s life.

3.3.2.9 Synthesis

We have now described all the types of nodes and relations that constitute a SIG. Figure 3.18 shows an overall SIG for “The Wily Lion” (except for the textual layer, which is given in Table 3.4, and $P_b$, which is omitted for brevity). State and Timeline nodes, as well as their related arcs, are not shown; as a notational convenience, we instead draw followed by arcs directly between $P$ nodes when the nodes are attached to subsequent State nodes. The interpretative layer here includes nodes representing many of the features we set out to model: the lion’s motivation (to eat the bull for purposes of enhancing his health), the problem blocking his goal (that the bull is dangerous due to his horns), his plan for overcoming the problem (to flatter the bull into forming a plan that involves removing his horns), his attempt at actualizing the plan (flattering compliments and pointed suggestions), and the successful outcome of the plan. The graph models the notion that the bull is the net loser in the transaction, having lost its life in an attempt to provide for its ego. A series of arcs connect textbase propositions to these meaning structures, either because they explicitly state the content of the structures, imply (entail) the content, or can be otherwise inferred to mean that the content is actualized or ceased. The textbase propositions, in turn, are mapped to discourse utterances in Table 3.4; though these interpreted as arcs are not drawn in Figure 3.18 due to space constraints, this mapping explicates the relationship between story time and telling time. The result is an encoding of a theory-of-mind interpretation of the fable that integrates several aspects of narrative, including agency and time, without relying on a prescriptive model of discourse structure such as a grammar.

3.3.3 Summary and Comparison to Prior Work

In this section we have described the 8 types of nodes and 18 relations that constitute the schemata of a Semantic Intention Graph. Table 3.8 summarizes the model in terms of the arc adjacencies that are defined for each possible pair of node types, with rows representing originating nodes and columns representing destination nodes. For instance, a Goal frame can relate to a Belief frame with in, would cause, would prevent, precondition for and precondition against. A + after an arc type indicates that more than one outgoing arc for the type is legal; otherwise, at most one outgoing arc is permitted.

We see the SIG as a next step in the evolution of narrative discourse models. For comparison, we have given four diagrams of “The Wily Lion” using different representations: Trabasso’s causal network formalism (Figure 3.2), Mandler and Johnson’s grammar (Figure 3.3), Lehnert’s plot units (Figure 3.5) and finally the SIG (Table 3.4 and Figure 3.18). The features of the SIG overlap with those of each model, but as a whole, it is a novel approach to diagramming narratives.

Table 3.9 outlines the way the four models differ in terms of their purpose and design. The interpretative layer of the SIG resembles plot units: In both cases, one can enumerate a
Figure 3.18: Overall encoding for “The Wily Lion” (textual layer shown in Table 3.4).
set of small, canonical “subgraphs” that represent thematic elements (such as loss) and chain them together to form arbitrarily large and complex structures. However, SIGs address many of the shortcomings of plot units by including a representation of time, a connection to the original discourse, and a more expressive representation of outcomes and mental states. Like GRTNs and grammars, SIGs show the connections between the functional components of a discourse—diagrams such as Figure 3.18 can connect any two timeline propositions, and by extension discourse clauses, that both relate to the same interpretative node (either directly or through closure rules). However, the SIG is more expressive than either of these models, as each imposes a rigid structure that excludes what we would consider to be thematically rich stories. For instance, it is difficult in either case to model two-agent interactions where mutual beliefs are important; agency frames allow us to represent such details. Overall, the SIG captures more thematic content in a narrative discourse than any of these four models, with respect to a theory-of-mind reading of a text.

### 3.4 Conclusion

This chapter has introduced the Story Intention Graph (SIG) as a closed set of discourse relations that collectively represent aspects of an agentive reading of narrative discourse. We reviewed four prior approaches and found them to have a limited expressive range, especially in representing elements of agency that research in cognition has shown to be key to narrative comprehension (agentive goals, plans, beliefs and attempts). Building on this prior work, we described the SIG schemata as emphasizing these and other facets that differentiate a story from an expository text or a set of disassociated facts. We do not claim that every discourse must be interpretable as a story, or that every story must feature easily discernible goals; indeed, one can find selections of modernist fiction that eschew both of these conventions. Rather, we claim that the SIG is an expressive, yet formal model for representing thematic content in the volumes of narrative discourse which employ the
CHAPTER 3. STORY INTENTION GRAPHS

<table>
<thead>
<tr>
<th>Purpose</th>
<th>GRTN</th>
<th>M/J Grammar</th>
<th>Plot Units</th>
<th>SIG</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cognitive modeling</td>
<td>Cognitive modeling, Discourse</td>
<td>AI understanding, story summarization</td>
<td>Discourse (corpus analysis; NL understanding)</td>
</tr>
<tr>
<td>Structure</td>
<td>Semantic network</td>
<td>Parse tree</td>
<td>Semantic network</td>
<td>Semantic network</td>
</tr>
<tr>
<td>Approach</td>
<td>Descriptive, but inflexible</td>
<td>Prescriptive</td>
<td>Descriptive</td>
<td>Descriptive</td>
</tr>
<tr>
<td>Input</td>
<td>Textbase propositions</td>
<td>Discourse units</td>
<td>Event model</td>
<td>Discourse units</td>
</tr>
<tr>
<td>Implied content</td>
<td>None (organizes textbase)</td>
<td>None</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Time</td>
<td>Propositions organized into temporal chains</td>
<td>No semantic model of time</td>
<td>Events organized into temporal chains</td>
<td>Interval-based timelines mapped to discourse units</td>
</tr>
<tr>
<td>SRL/WSD compatible</td>
<td>Yes (textbase propositions)</td>
<td>No</td>
<td>No</td>
<td>Yes (propositions, discourse input)</td>
</tr>
<tr>
<td>Goals</td>
<td>Explicit (G nodes)</td>
<td>Explicit (GOAL)</td>
<td>Explicit &amp; Implicit (M nodes)</td>
<td>Explicit &amp; Implicit (G frames)</td>
</tr>
<tr>
<td>Plans</td>
<td>Subgoals (G→G)</td>
<td>Subgoals (nested GOAL_PATHS)</td>
<td>“Motivation” units (M→m→M)</td>
<td>Network of subgoals (sec, wp, pf, pa)</td>
</tr>
<tr>
<td>Beliefs</td>
<td>Explicit (R nodes)</td>
<td>Explicit (Internal Event)</td>
<td>Explicit &amp; Implicit (M nodes)</td>
<td>Explicit &amp; Implicit (B frames)</td>
</tr>
<tr>
<td>Attempts</td>
<td>Explicit (A nodes)</td>
<td>Explicit (ATTEMPT)</td>
<td>None</td>
<td>Explicit &amp; Implicit (ac, ap)</td>
</tr>
<tr>
<td>Outcomes</td>
<td>Explicit (O nodes)</td>
<td>Explicit (OUTCOME)</td>
<td>Explicit &amp; Implicit (a →+, a →−)</td>
<td>Explicit &amp; Implicit (actualization status)</td>
</tr>
<tr>
<td>Affect</td>
<td>Explicit (SO vs. AO outcomes)</td>
<td>None</td>
<td>Complete (+ and −)</td>
<td>Complete (Affect nodes; goal closure)</td>
</tr>
<tr>
<td>Theory of Mind</td>
<td>Single-character POV</td>
<td>None</td>
<td>Multiple broad domains</td>
<td>Nestable agency frames</td>
</tr>
<tr>
<td>Implementation</td>
<td>Machine simulation; cognitive experiments</td>
<td>Cognitive experiments</td>
<td>AI system, cognitive experiments</td>
<td>Software platform &amp; annotation UI; collection project</td>
</tr>
</tbody>
</table>

Table 3.9: Comparison between Trabasso’s GRTN model, Mandler and Johnson’s story grammar model, Lehnert’s plot-unit model, and the SIG model.

Let us conclude this chapter by revisiting once more the citation from E. M. Forster about the difference between a non-story and a story:

1. The king died. Then the queen died.
2. The king died. Then the queen died of grief.

Forster’s point is that a set of timeline-ordered events is not necessarily a narrative. There must be other inter-sentential relations that bind the discourse together. Causality and motivation are the underpinnings of our approach, and in the case of (2), they relate the second sentence to the first.

Figure 3.19 illustrates this effect graphically. SIG encodings are drawn for both (1) and (2) in 3.19(i) and 3.19(ii), respectively. In both cases, the textual layer contains the original...
discourse clauses, and the timeline layer dutifully constructs a timeline of sequential events. But only the second version provides a rationale for relating the two deaths as a causally coherent whole, and the interpretative layer provides the means to represent the difference.

On a broader scale, we may also consider how SIGs rate according to the criteria for a new representation we defined at the beginning of this chapter: expressiveness, robustness, computability and accessibility.

First, is the SIG schemata expressive enough to have wide coverage over the range of what a reasonable reader would consider to be thematically rich narratives? We believe that SIGs are highly expressive due to the “open-ended” nature of the graph architecture. We have defined a set of relations that can be instantiated and combined in patterns to reflect an extensible range of thematic content, just as a closed set of words in a lexicon can be combined in syntactic patterns to form a far larger set of possible sentences. The relations serve as common building blocks for a range of narrative situations congruent with a theory-of-mind reading of a text (such as revenge, deception, success, failure, and regret) as well as formal storytelling devices (flashbacks, point of view, mystery, and so on). Because new stories are constantly being told, recombined as they may be from previously told stories, we cannot prove by exhaustion that the SIG is sufficiently expressive to cover every possible discourse that a reasonable reader would consider narrative in nature. However, we can show that the SIG is more expressive than previous descriptive models we have considered, and by enumerating a set of SIG fragments that model a wide-ranging set of narrative scenarios, we can demonstrate the framework of a wide expressive range by example. We take up this task at length in Appendix B.

Second, is a SIG robust, so that it gracefully handles varying degrees of semantic precision? Yes—as we have designed it in such a way that partial encodings of a story are permissible. Being a descriptive formalism, the SIG allows multiple levels of abstraction.
There is no prescribed number of nodes that must be instantiated, though we will discuss guidelines for human annotation in Chapter 4. One can, for instance, forego a detailed discussion of a multi-step plan and label all of the actions by an agent as an “attempt to cause” a positive or negative affect state. (The wily lion, in such a flat reading, did everything in an attempt to fulfill a one-step plan to improve his health.) The formalism is also agnostic to the type of representation associated with each Proposition (P), Interpretative Proposition (I) and Affect (A) node—one can combine the SIG relations with any type of sentential knowledge representation (such as propositions) using I and P nodes as containers, or assign nothing to them except for agent metadata. In the latter case, the relations still describe the thematic aspects of a highly abstract story, such as one about overcoming adversity:

“Agent X wanted to do A because he thought it would help him do B. Agent Y tried to prevent X from doing A. In the end, Agent Y prevented Agent X from doing A. But Agent X did B through other means.”

This aspect of the SIG also gives it a domain independence, in that no particular predicate vocabulary is defined to be part of the model; although we focus on Aesop’s fables for their brevity, we will soon apply the model to other genres and longer stories. We take up the question of whether the schemata is computable in Chapter 5.

The final question is that of the accessibility of our approach with respect to trained annotators. We will explore this in the following chapter, in which we implement a software platform and annotation interface for creating a DramaBank of story encodings.