Tense and Aspect Assignment in Narrative Discourse

David K. Elson and **Kathleen R. McKeown**
Department of Computer Science
Columbia University
{delson,kathy}@cs.columbia.edu

Abstract

We describe a method for assigning English tense and aspect in a system that realizes surface text for symbolically encoded narratives. Our testbed is an encoding interface in which propositions that are attached to a timeline must be realized from several temporal viewpoints. This involves a mapping from a semantic encoding of time to a set of tense/aspect permutations. The encoding tool realizes each permutation to give a readable, precise description of the narrative so that users can check whether they have correctly encoded actions and statives in the formal representation. Our method selects tenses and aspects for individual event intervals as well as subintervals (with multiple reference points), quoted and unquoted speech (which reassign the temporal focus), and modal events such as conditionals.

1 Introduction

Generation systems that communicate knowledge about time must select tense and aspect carefully in their surface realizations. An incorrect assignment can give the erroneous impression that a continuous action has ended, or that a previous state is the current reality. In this paper, we consider English tense and aspect in the generation of narrative discourse, where statives and actions occur over connected intervals.

We describe two contributions: first, a general application of theories of tense, aspect and interval logic to a generation context in which we map temporal relationships to specific tense/aspect selections. Second, we describe an implementation of this approach in an interactive environment with a basic sentence planner and realizer. The first result does not depend on the second.

The purpose of the system is to allow users who are naïve to linguistics and knowledge representation to create semantic encodings of short stories. To do this, they construct propositions (predicate-argument structures) through a graphical, menu-based interface, and assign them to intervals on a timeline. Figure 1 shows a session in which the user is encoding a fable of Aesop. The top-right panel shows the original fable, and the left-hand panel shows a graphical timeline with buttons for constructing new propositions at certain intervals. The left-hand and bottom-right panels contain automatically generated text of the encoded story, as the system understands it, from different points of view. Users rely on these realizations to check that they have assigned the formal connections correctly. The tenses and aspects of these sentences are a key component of this feedback. We describe the general purpose of the system, its data model, and the encoding methodology in a separate paper (Elson and McKeown, 2010).

The paper is organized as follows: After discussing related work in Section 2, we describe our method for selecting tense and aspect for single events in Section 3. Section 4 follows with more complex cases involving multiple events and shifts in temporal focus. We then discuss the results.

2 Related Work

There has been intense interest in the interpretation of tense and aspect into a formal understanding of the ordering and duration of events. This work has been in both linguistics (Dowty, 1979; Nerbonne, 1986; Vlach, 1993) and natural language understanding. Early systems investigated rule-based approaches to parsing the durations and orderings of events from the tenses and aspects of their verbs (Hinrichs, 1987; Webber, 1987; Song and Cohen, 1988; Passonneau, 1988). Allen (1984) and Steedman (1995) focus on distinguishing between achievements (when an event culminates in a result, such as *John builds a house*) and processes (such as walking). More

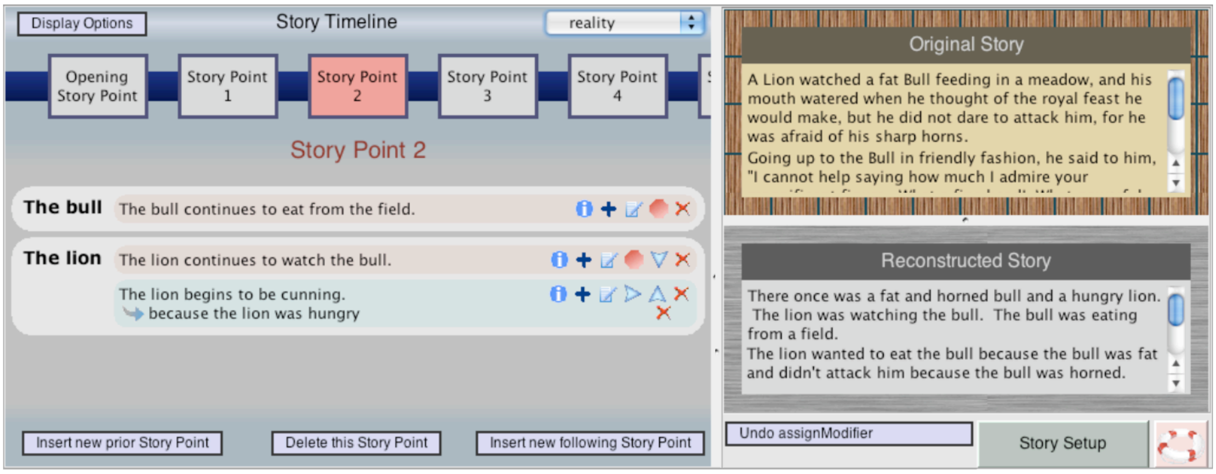


Figure 1: Screenshot of our story encoding interface.

recent work has centered on markup languages for complex temporal information (Mani, 2004) and corpus-based (statistical) models for predicting temporal relationships on unseen text (Mani et al., 2006; Lapata and Lascarides, 2006).

Our annotation interface requires a fast realizer that can be easily integrated into an interactive, online encoding tool. We found that developing a custom realizer as a module to our Java-based system was preferable to integrating a large, general purpose system such as KPML/Nigel (Matthiessen and Bateman, 1991) or FUF/SURGE (Elhadad and Robin, 1996). These realizers, along with RealPro (Lavoie and Rambow, 1997), accept tense as a parameter, but do not calculate it from a semantic representation of overlapping time intervals such as ours (though the Nigel grammar can calculate tense from speech, event, and reference time orderings, discussed below). The statistically trained FERGUS (Chen et al., 2002) contrasts with our rule-based approach.

Dorr and Gaasterland (1995) and Grote (1998) focus on generating temporal connectives, such as *before*, based on the relative times and durations of two events; Gagnon and Lapalme (1996) focus on temporal adverbials (e.g., when to insert a known time of day for an event). By comparison, we extend our approach to cover direct/indirect speech and the subjunctive/conditional forms, which they do not report implementing. While our work focuses on English, Yang and Bateman (2009) describe a recent system for generating Chinese aspect expressions based on a time interval representation, using KPML as their surface realizer.

Several other projects run tangential to our interactive narrative encoding project. Callaway and Lester's STORYBOOK (2002) aims to improve fluency and discourse cohesion in realizing formally encoded narratives; Ligozat and Zock (1992) allow users to interactively construct sentences in various temporal scenarios through a graphical interface.

3 Expressing single events

3.1 Temporal knowledge

The propositions that we aim to realize take the form of a predicate, one or more arguments, zero or more attached modifiers (either a negation operator or an adverbial, which is itself a proposition), and an assignment in time. Each argument is associated with a semantic role (such as Agent or Experiencer), and may include nouns (such as characters) or other propositions. In our implemented system, the set of predicates available to the annotator is adapted from the VerbNet (Kingsbury and Palmer, 2002) and WordNet (Fellbaum, 1998) linguistic databanks. These provide both durative actions and statives (Dowty, 1979); we will refer to both as *events* as they occur over intervals. For example, here are an action and a stative:

$$\texttt{walk}(\texttt{Mary}, \texttt{store}, 2, 6) \quad (1)$$

$$\texttt{hungry}(\texttt{Julia}, 1, \infty) \quad (2)$$

The latter two arguments in (1) refer to time states in a totally ordered sequence; Mary starts walking to the store at state 2 and finishes walking at state 6. (2) begins at state 1, but is unbounded (Julia never ceases being hungry). While this paper does not address the use of reference times (such as equating a state to 6:00 or *yesterday*), this is an area of ongoing work.

(1) and (2), depending on the situation, can be realized in several aspects and tenses. We adapt and extend Reichenbach's (1947) famous system of symbols for distinguishing between simple and progressive aspect. Reichenbach identifies three points that define the temporal position of the event: the event time E, the speech time S, and a reference time R which may or may not be indicated by a temporal adverbial. The total ordering between these times dictates the appropriate aspect. For example, the simple past *John laughed* has the relation $E < S$. $R = E$ because there is no separate reference time involved. The past perfect *John had laughed [by the end of the play]* has the relation $E < R < S$, in that it describe "the past of the past", with the nearer "past" being R (the end of the play). R can be seen as the temporal focus of the sentence.

As Reichenbach does not address events with intervals, we redefine E as the transition $(E_1..E_2)$ attached to the proposition (for example, (2,6) for Mary's walk). This definition deliberately assumes that no event ever occurs over a single "instant" of time. The perception of an instantaneous event, when it is needed, is instead created by dilating R into an interval large enough to contain the entire event, as in Dowty (1979).

We also distinguish between two generation modes: realizing the story as a complete discourse (*narration mode*) and describing the content of a single state or interval (*snapshot mode*). Our system supports both modes differently. In narration mode, we realize the story as if all events occur before the speech time S, which is the style of most literary fiction. (We shall see that this does not preclude the use of the future tense.) In snapshot mode, speech time is concurrent with reference time so that the same events are realized as though they are happening "now." The system uses this mode to allow annotators to inspect and edit what occurs at any point in the story. In Figure 1, for instance, the lion's watching of the bull is realized as both a present, continuing event in snapshot mode (*the lion continues to watch the bull*) and narrated as a past, continuing event (*the lion was watching the bull*). In both cases, we aim to precisely translate the propositions and their temporal relationships into text, even if the results are not elegant rhetoric, so that annotators can see how they have formally encoded the story. In the remainder of this section, we describe our method for assigning tenses and aspects to propositions such as these.

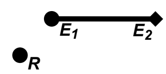
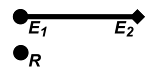
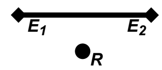
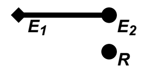
Diagram	Relations	Perspective
	$R < E_1$	Before
	$R = E_1$ $R < E_2$	Begin
	$E_1 < R$ $R < E_2$	During
	$R = E_2$ $R > E_1$	Finish
	$R > E_2$	After

Table 1: Perspective assignment for viewing an event from a reference state.

3.2 Reference state

In both snapshot and narration modes, we often need to render the events that occur at some reference state R. We would like to know, for instance, what is happening now, or what happened at 6:00 yesterday evening. The tense and aspect depend on the *perspective* of the reference state on the event, which can be bounded or unbounded. The two-step process for this scenario is to determine the correct perspective, then pick the tense and aspect class that best communicates it.

We define the set of possible perspectives to follow Allen (1983), who describes seven relationships between two intervals: before/after, meets/met by, overlaps/overlapped by, starts/started by, during/contains, finishes/finished by, and equals. Not all of these map to a relationship between a single reference *point* and an event interval. Table 1 maps each possible interaction between E and R to a perspective, for both bounded and unbounded events, including the defining relationships for each interaction. A diamond for E_1 indicates *at or before*, i.e., the event is either anteriorly unbounded ($E_1 = -\infty$) or beginning at a state prior to R and E_2. Similarly, a diamond for E_2 indicates *at or after*.

Once the perspective is determined, covering Reichenbach's E and R, speech time S is determined by the generation mode. Following the guidelines of Reichenbach and Dowty, we then assign a tense for each perspective/speech time per-

Perspective	Generation mode	**English tense**	System's construction	Example
After	Future Speech	Past perfect	*had* {PAST PARTICIPLE}	She had walked.
	Present Speech	Present perfect	*has/have* {PAST PARTICIPLE}	She has walked.
	Past Speech	Future perfect	*will have* {PAST PARTICIPLE}	She will have walked.
	Modal Infinitive		*to have* {PAST PARTICIPLE}	To have walked.
Finish	Future Speech	"Finished"	*stopped* {PROGRESSIVE}	She stopped walking.
	Present Speech	"Finishes"	*stops* {PROGRESSIVE}	She stops walking.
	Past Speech	"Will finish"	*will stop* {PROGRESSIVE}	She will stop walking.
	Modal Infinitive		*to stop* {PROGRESSIVE}	To stop walking.
During	Future Speech	Past progressive	*was/were* {PROGRESSIVE}	She was walking.
	Present Speech	Present progressive	*am/is/are* {PROGRESSIVE}	She is walking.
	Past Speech	Future progressive	*will be* {PROGRESSIVE}	She will be walking.
	Modal Infinitive		*to be* {PROGRESSIVE}	To be walking.
During-After	Future Speech	Past perfect progressive	*had been* {PROGRESSIVE}	She had been walking.
	Present Speech	Present perfect progressive	*has/have been* {PROGRESSIVE}	She has been walking.
	Past Speech	Future perfect progressive	*will have been* {PROGRESSIVE}	She will have been walking.
	Modal Infinitive		*to has/have been* {PROGRESSIVE}	To have been walking.
Begin	Future Speech	"Began"	*began* {INFINITIVE}	She began to walk.
	Present Speech	"Begins"	*begins* {INFINITIVE}	She begins to walk.
	Past Speech	"Will begin"	*will begin* {INFINITIVE}	She will begin to walk.
	Modal Infinitive		*to begin* {PROGRESSIVE}	To begin walking.
Contains	Future Speech	Simple past	{SIMPLE PAST}	She walked.
	Present Speech	Simple present	{SIMPLE PRESENT}	She walks.
	Past speech	Simple future	*will* {INFINITIVE}	She will walk.
	Modal Infinitive		{INFINITIVE}	To walk.
Before	Future Speech	"Posterior"	*was/were going* {INFINITIVE}	She was going to walk.
	Present Speech	Future	*am/is/are going* {INFINITIVE}	She is going to walk.
	Past Speech	Future-of-future	*will be going* {INFINITIVE}	She will be going to walk.
	Modal Infinitive		*to be going* {INFINITIVE}	To be going to walk.

Table 2: Tense/aspect assignment and realizer constructions for describing an action event from a particular perspective and speech time. "Progressive" means "present participle."

mutation in Table 2. Not all permutations map to actual English tenses. Narration mode is shown as *Future Speech*, in that S is in the future with respect to all events in the timeline. (This is the case even if E is unbounded, with $E_2 = \infty$.) Snapshot mode is realized as *Present Speech*, in that $R = S$. The fourth column indicates the syntactic construction with which our system realizes the permutation. Each is a sequence of tokens that are either cue words (*began, stopped*, etc.) or conjugations of the predicate's verb. These constructions emphasize precision over fluency.

As we have noted, theorists have distinguished between "statives" that are descriptive (*John was hungry*), "achievement" actions that culminate in a state change (*John built the house*), and "activities" that are more continuous and divisible (*John read a book for an hour*) (Dowty, 1979). Prior work in temporal connectives has taken advantage of lexical information to determine the correct situation and assign aspect appropriately (Moens and Steedman, 1988; Dorr and Gaasterland, 1995). In our case, we only distinguish between actions and statives, based on information from WordNet and VerbNet. We use a separate table for statives; it is similar to Table 2, except the constructions replace verb conjugations with insertions of *be*, *been*, *being*, *was*, *were*, *felt*, and so on (with the latter applying to affective states). We do not currently distinguish between achievements and activities in selecting tense and aspect, except that the annotator is tasked with "manually" indicating a new state when an event culminates in one (e.g., *The house was complete*). Recognizing an achievement action can benefit lexical choice (better to say *John finished building the house* than *John stopped*) and content selection for the discourse as a whole (the house's completion is implied by *finished* and does not need to be stated separately).

To continue our running examples, suppose propositions (1) and (2) were viewed as a snapshot from state $R = 2$. Table 1 indicates *Begin*

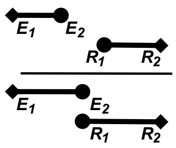
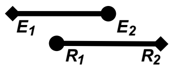
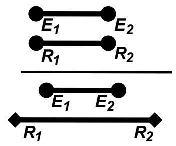
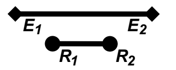
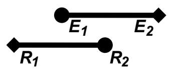
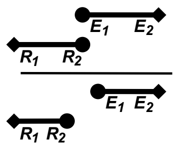
Diagram	Relations	Perspective
	$R_1 \geq E_2$	After
	$R_1 > E_1$ $E_2 > R_1$ $R_2 > E_2$	Finish
	$R_1 \leq E_1$ $R_2 \geq E_2$	Contains
	$E_1 < R_1$ $E_2 > R_2$	During
	$R_1 < E_1$ $R_2 > E_1$ $E_2 > R_2$	Begin
	$E_1 \geq R_2$	Before

Table 3: Perspective assignment for describing an event from an assigned perspective.

to be the perspective for (1), since $E_1 = R$, and Table 2 calls for a "new" tense/aspect permutation that means "begins at the present time." When the appropriate construction is inserted into the overall syntax for *walk(Agent, Destination)*, which we derive from the VerbNet frame for *walk*, the result is *Mary begins to walk to the store*; similarly, (2) is realized as *Julia is hungry* via the *During* perspective. Narration mode invokes past-tense verbs.

3.3 Reference interval

Just as events occur over intervals, rather than single points, so too can reference times. One may need to express what occurred when "Julia entered the room" (a non-instantaneous action) or "yesterday evening." Our system allows annotators to view intervals in snapshot mode to get a sense of what happens over a certain time span.

The semantics of reference intervals have been studied as extensions to Reichenbach's point approach. Dowty (1979, p.152), for example, posits that the progressive fits only if the reference interval is completely contained within the event interval. Following this, we construct an alternate lookup table (Table 3) for assigning the perspective of an event from a reference interval. Table 2 then applies in the same manner. In snapshot mode, the speech time S also occurs over an interval (namely, R), and Present Speech is still used. In narration mode, S is assumed to be a point following all event and reference intervals. In our running example, narrating the interval (1,7) results in *Mary walked to the store* and *Julia began to be hungry*, using the *Contains* and *Begin* perspectives respectively.

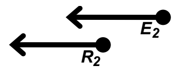
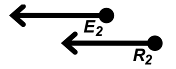
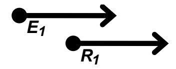
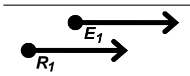
Diagram	Relations	Perspective
	$E_2 > R_2$ $E_1 = -\infty$ $R_1 = -\infty$	During (*a priori*)
	$R_2 > E_2$ $E_1 = -\infty$ $R_1 = -\infty$	After
	$R_1 > E_1$ $E_2 = \infty$ $R_2 = \infty$	Contains
	$E_1 > R_1$ $E_2 = \infty$ $R_2 = \infty$	Before

Table 4: Perspective assignment if event and reference intervals are unbounded in like directions.

The notion of an *unbounded* reference interval, while unusual, corresponds to a typical perspective if the event is either bounded or unbounded in the opposite direction. These scenarios are illustrated in Table 3. Less intuitive are the cases where event and reference intervals are unbounded in the same direction. Perspective assignments for these instances are described in Table 4 and emphasize the bounded end of R. These situations occur rarely in this generation context.

3.4 Event Subintervals

We do not always want to refer to events in their entirety. We may instead wish to refer to the beginning, middle or end of an event, no matter when it occurs with respect to the reference time. This invokes a second reference point in the same interval (Comrie, 1985, p.128), delimiting a subinterval. Consider *John searches for his glasses* versus *John continues to search for his glasses*– both indicate an ongoing process, but the latter implies a subinterval during which time, we are expected to know, John was already looking for his glasses.

Our handling of subintervals falls along four alternatives that depend on the interval $E_1..E_2$, the reference R and the subinterval $E'_1..E'_2$ of E, where $E'_1 \geq E_1$ and $E'_2 \leq E_2$.

1. **During-After.** If E' is not a final subinterval of E ($E'_2 < E_2$), and $R = E'_2$ or R is a subinterval of E that is met by E' ($R_1 = E'_2$), the perspective of E' is defined as *During-After*. In Table 2, this invokes the perfect-progressive tense. For example, viewing example (1) with $E' = (2, 4)$ from $R = 4$ in narration mode (Future Speech) would yield *Mary had been walking to the store.*

2. **Start.** Otherwise, if E' is an initial subinterval of E ($E'_1 = E_1$ and $E'_2 < E_2$), the perspective is defined as *Start*. These rows are omitted from Table 2 for space reasons, but the construction for this case reassigns the perspective to that between R and E'. Our realizer reassigns the verb predicate to *begin* (or *become* for statives) with a plan to render its only argument, the original proposition, in the infinitive tense. For example, narrating (2) with E' =(1,2) from $R = 3$ would yield *Julia had become hungry*.

3. **Continue.** Otherwise, and similarly, if E strictly contains E' ($E'_1 > E_1$ and $E'_2 < E_2$), we assign the perspective *Continue*. To realize this, we reassign the perspective to that between R and E', and reassign the verb predicate to *continue* (or *was still* for statives) with a plan to render its only argument, the original proposition, in the infinitive.

4. **End.** Otherwise, if E' is a final subinterval of E ($E'_1 > E_1$ and $E'_2 = E_2$), we assign the perspective *End*. To realize this, we reassign the perspective to that between R and E', and reassign the verb predicate to *stop* (or *finish* for cumulative achievements). Similarly, the predicate's argument is the original proposition rendered in the infinitive.

4 Alternate timelines and modalities

This section covers more complex situations involving *alternate timelines*– the feature of our representation by which a proposition in the main timeline can refer to a second frame of time. Other models of time have supported similar encapsulations (Crouch and Pulman, 1993; Mani and Pustejovsky, 2004). The alternate timeline can contain references to actual events or modal events (imagined, obligated, desired, planned, etc.) in the past the future with respect to its *point of attachment* on

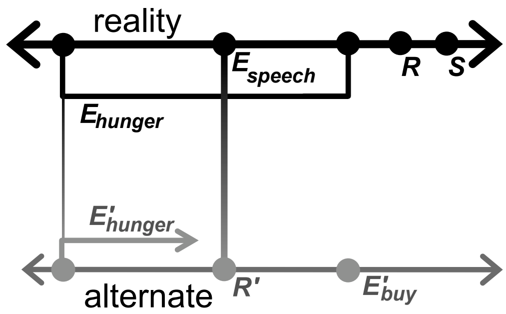


Figure 2: Schematic of a speech act attaching to a alternate timeline with a hypothetical action. R' and E_{speech} are attachment points.

the main timeline. This is primarily used in practice for modeling dialogue acts, but it can also be used to place real events at uncertain time states in the past (e.g., the present perfect is used in a reference story being encoded).

4.1 Reassigning Temporal Focus

Ogihara (1995) describes dialogue acts involving changes in temporal focus as "double-access sentences." We now consider a method for planning such sentences in such a way that the refocusing of time (the reassignment of R into a new context) is clear, even if it means changing tense and aspect mid-sentence. Suppose Mary were to declare that she would buy some eggs because of Julia's hunger, but before she returned from the store, Julia filled up on snacks. If this speech act is described by a character later in the story, then we need to carefully separate what is known to Mary at the time of her speech from what is later known at R by the teller of the episode. Mary sees her purchase of eggs as a possible future, even though it may have already happened by the point of retelling, and Mary does not know that Julia's hunger is to end before long.

Following Hornstein's treatment of these scenarios (Hornstein, 1990), we attach R', the reference time for Mary's statement (in an alternate timeline), to E_{speech}, the event of her speaking (in the main timeline). The act of buying eggs is a hypothetical event E'_{buy} that falls after R' on the alternate (modal) timeline. S is not reassigned.

Figure 2 shows both timelines for this example. The main timeline is shown on top; Mary's speech act is below. The attachment point on the main timeline is, in this case, the speech event E_{speech}; the attachment point on an alternate timeline is always R'. The placement of R, the main reference point, is not affected by the alternate timeline. Real events, such as Julia's hunger, can be invoked in the alternate timeline (as drawn with a vertical line from E_{hunger} to an E'_{hunger} without an E'_2 known to Mary) but they must preserve their order from the main timeline.

The tense assignment for the event intervals in the alternate timeline then proceeds as normal, with R' substituting for R. The hypothetical "buy" event is seen in *Before* perspective, but past tense (Future Speech), giving the "posterior" (future-of-a-past) tense. Julia's hunger is seen as *During* as per Table 1. Further, we assert that connectives such as *Because* do not alter R (or in this situation, R'), and that the E'_{buy} is connected to E'_{hunger} with a causality edge. (Annotators can indicate connectives between events for causality, motivation and other features of narrative cohesion.)

The result is: *Mary had said that she was going to buy eggs because Julia was hungry.* The subordinate clause following *that* sees E'_{buy} in the future, and E'_{hunger} as ongoing rather than in the past. It is appropriately ambiguous in both the symbolic and rendered forms whether E'_{buy} occurs at all, and if so, whether it occurs before, during or after R. A discourse planner would have the responsibility of pointing out Mary's mistaken assumption about the duration of Julia's hunger.

We assign tense and aspect for quoted speech differently than for unquoted speech. Instead of holding S fixed, S' is assigned to R' at the attachment point of the alternate timeline (the "present time" for the speech act). If future hypothetical events are present, they invoke the Past Speech constructions in Table 2 that have not been used by either narration or snapshot mode. The content of the quoted speech then operates totally independently of the speech action, since both R' and S' are detached: *Mary said/says/was saying, "I am going to buy eggs because Julia is hungry."*

The focus of the sentence can be subsequently reassigned to deeper nested timelines as necessary (attaching E' to R'', and so on). Although the above example uses subordinate clauses, we can use this nesting technique to construct composite tenses such as those enumerated by Halliday (1976). To this end, we conjugate the *Modal Infinitive* construction in Table 2 for each alternate timeline. For instance, Halliday's complex form "present in past in future in past in future" (as in *will have been going to have been taking*) can be generated with four timelines in a chain that invoke, in order and with Past Speech, the *After*, *Before*, *After* and *During* perspectives. There are four Rs, all but the main one attached to a previous E.

4.2 Subjunctives and Conditionals

We finally consider tense and aspect in the case of subjunctive and conditional statements (if-thens), which can appear in alternate timelines. The relationship between an *if* clause and a *then* clause is not the same as the relationship between two clauses joined by *because* or *when*. The *then* clause– or set of clauses– is predicated on the truth of the *if* clause. As linguists have noted (Hornstein, 1990, p.74), the *if* clause serves as an adverbial modifier, which has the effect of moving forward the reference point to the last of the *if* event intervals (provided that the *if* refers to a hypothetical future). Consider the sentence: *If John were to fly to Tokyo, he would have booked a hotel.* A correct model would place E'_{book} before E'_{fly} on an alternate timeline, with E'_{fly} as the *if*. Since *were to fly* is a hypothetical future, $R' < E'_{fly}$. During regeneration, we set R' to E'_{fly} after rendering *If John were to fly to Tokyo*, because we begin to assume that this event transpired. If R' is left unchanged, it may be erroneously left before E'_{book}: *Then he would be going to book a hotel.*

Our encoding interface allows users to mark one or more events in an alternate timeline as *if* events. If at least one event is marked, all *if* events are rendered in the subjunctive mood, and the remainder are rendered in the conditional. For the *if* clauses that follow R', S' and R' itself are reassigned to the interval for each clause in turn. R' and S' then remain at the latest *if* interval (if it is after the original R') for purposes of rendering the *then* clauses. In our surface realizer, auxiliary words *were* and *would* are combined with the Modal Infinitive constructions in Table 2 for events during or following the original attachment point.

As an example, consider an alternate timeline with two statives whose start and end points are the same: *Julia is hungry* and *Julia is unhappy*. The former is marked *if*. Semantically, we are saying that `hungry(Julia)→unhappy(Julia)`. If R' were within these intervals, the rendering would be: *If Julia is hungry, then she is unhappy* (*Contains*/Present Speech for both clauses). If R' were prior to these intervals, the rendering would be: *If Julia were to be hungry, then she would be unhappy*. This reassigns R' to E_{hungry}, using *were* as a futurative and *would* to indicate a conditional. Because R' and S' are set to E_{hungry}, the perspective on both clauses remains *Contains*/Present Speech. Finally, if both intervals are before R', describing Julia's previous emotional states, we avoid shifting R' and S' backward: *If Julia had been hungry, then she had been unhappy* (*After* perspective, Future Speech for both statives).

The algorithm is the same for event intervals. Take (1) and a prior event where Mary runs out of eggs:

$$\texttt{runOut}(\texttt{Mary}, \texttt{eggs}, 0, 1) \quad (3)$$

Suppose they are in an alternate timeline with attachment point $0'$ and (1) marked *if*. We begin by realizing Mary's walk as an *if* clause: *If Mary were to walk to the store*. We reassign R' to E_{walk}, (2,6), which diverts the perception of (3) from *Begins* to *After*: *She would have run out of eggs*. Conversely, suppose the conditional relationship were reversed, with (3) as the only *if* action. If the attachment point is $3'$, we realize (3) first in the *After* perspective, as R' does not shift backward: *If Mary had run out of eggs*. The remainder is rendered from the *During* perspective: *She would be walking to the store*. Note that in casual conversation, we might expect a speaker at $R = 3$ to use the past simple: *If Mary ran out of eggs, she would be walking to the store*. In this case, the speaker is attaching the alternate timeline at a reference interval that subsumes (3), invoking the *Contains* perspective by casting a net around the past. We ask our annotators to select the best attachment point manually; automatically making this choice is beyond the scope of this paper.

5 Discussion

As we mentioned earlier, we are describing two separate methods with a modular relationship to one another. The first is an abstract mapping from a conceptual representation of time in a narrative, including interval and modal logic, to a set of 11 *perspectives*, including the 7 listed in Table 2 and the 4 introduced in Section 3.4. These 11 are crossed with three scenarios for speech time to give a total of 33 tense/aspect permutations. We also use an infinitive form for each perspective. One may take these results and map them from other time representations with similar specifications.

The second result is a set of syntactic constructions for realizing these permutations in our story encoding interface. Our focus here, as we have noted, is not fluency, but a surface-level rendering that reflects the relationships (and, at times, the ambiguities) present in the conceptual encoding. We consider variations in modality, such as an indicative reading as opposed to a conditional or subjunctive reading, to be at the level of the realizer and not another class of tenses.

We have run a collection project with our encoding interface and can report success in the tool's usability (Elson and McKeown, 2009). Two annotators each encoded 20 fables into the formal representation, with their only exposure to the semantic encodings being through the reference text generator (as in Figure 1). Both annotators became comfortable with the tool after a period of training; in surveys that they completed after each task, they gave Likert-scale usability scores of 4.25 and 4.30 (averaged over 20 tasks, with 5 meaning "easiest to use"). These scores are not specific to the generation component, but they suggest that annotators could derive satisfactory tenses from their semantic structures. The most frequently cited deficiency in the model in terms of time was the inability to assign reference times to states and intervals (such as *the next morning*).

6 Conclusion and Future Work

It has always been the goal in surface realization to generate sentences from a purely semantic representation. Our approach to the generation of tense and aspect from temporal intervals takes us closer to that goal. We have applied prior work in linguistics and interval theory and tested our approach in an interactive narrative encoding tool. Our method handles reference intervals and event intervals, bounded and unbounded, and extends into subintervals, modal events, conditionals, and direct and indirect speech where the temporal focus shifts.

In the future, we will investigate extensions to the current model, including temporal adverbials (which explain the relationship between two events), reference times, habitual events, achievements, and discourse-level issues such as preventing ambiguity as to whether adjacent sentences occur sequentially (Nerbonne, 1986; Vlach, 1993).

7 Acknowledgments

This material is based on research supported in part by the U.S. National Science Foundation (NSF) under IIS-0935360. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

References

James F. Allen. 1983. Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26(11):832–843.

James F. Allen. 1984. Towards a general theory of action and time. *Artificial Intelligence*, 23(2):123–154.

Charles Callaway and James Lester. 2002. Narrative prose generation. *Artificial Intelligence*, 139(2):213–252.

John Chen, Srinivas Bangalore, Owen Rambow, and Marilyn Walker. 2002. Towards automatic generation of natural language generation systems. In *Proceedings of the 19th International Conference on Computational Linguistics (COLING 2002)*, Taipei, Taiwan.

Bernard Comrie. 1985. *Tense*. Cambridge University Press.

Richard Crouch and Stephen Pulman. 1993. Time and modality in a natural language interface to a planning system. *Artificial Intelligence*, pages 265–304.

Bonnie J. Dorr and Terry Gaasterland. 1995. Selecting tense, aspect, and connecting words in language generation. In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (IJCAI-95)*, Montreal, Canada.

David R. Dowty. 1979. *Word Meaning and Montague Grammar*. D. Reidel, Dordrecht.

Michael Elhadad and Jacques Robin. 1996. An overview of surge: a reusable comprehensive syntactic realization component. In *INLG '96 Demonstrations and Posters*, pages 1–4, Brighton, UK. Eighth International Natural Language Generation Workshop.

David K. Elson and Kathleen R. McKeown. 2009. A tool for deep semantic encoding of narrative texts. In *Proceedings of the ACL-IJCNLP 2009 Software Demonstrations*, pages 9–12, Suntec, Singapore.

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

Christiane Fellbaum. 1998. *WordNet: An Electronic Lexical Database*. MIT Press, Cambridge, MA.

Michel Gagnon and Guy Lapalme. 1996. From conceptual time to linguistic time. *Computational Linguistics*, 22(1):91–127.

Brigitte Grote. 1998. Representing temporal discourse markers for generation purposes. In *Proceedings of the Discourse Relations and Discourse Markers Workshop*, pages 22–28, Montreal, Canada.

M.A.K. Halliday. 1976. The english verbal group. In G. R. Kress, editor, *Halliday: System and Function in Language*. Oxford University Press, London.

Erhard W. Hinrichs. 1987. A compositional semantics of temporal expressions in english. In *Proceedings of the 25th Annual Conference of the Association for Computational Linguistics (ACL-87)*, Stanford, CA.

Norbert Hornstein. 1990. *As Time Goes By: Tense and Universal Grammar*. MIT Press, Cambridge, MA.

Paul Kingsbury and Martha Palmer. 2002. From treebank to propbank. In *Proceedings of the Third International Conference on Language Resources and Evaluation (LREC-02)*, Canary Islands, Spain.

Mirella Lapata and Alex Lascarides. 2006. Learning sentence-internal temporal relations. *Journal of Artificial Intelligence Research*, 27:85–117.

Benoit Lavoie and Owen Rambow. 1997. A fast and portable realizer for text generation systems. In *Proceedings of the Fifth Conference on Applied Natural Language Processing*, Washington, DC.

Gerard Ligozat and Michael Zock. 1992. How to visualize time, tense and aspect? In *Proceedings of the 14th International Conference on Computational Linguistics (COLING '92)*, pages 475–482, Nantes, France.

Inderjeet Mani and James Pustejovsky. 2004. Temporal discourse models for narrative structure. In *Proceedings of the ACL Workshop on Discourse Annotation*, Barcelona, Spain.

Inderjeet Mani, Marc Verhagen, Ben Wellner, Chong Min Lee, and James Pustejovsky. 2006. Machine learning of temporal relations. In *Proceedings of COLING/ACL 2006*, pages 753–760, Sydney, Australia.

Inderjeet Mani. 2004. Recent developments in temporal information extraction. In *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP '03)*, pages 45–60, Borovets, Bulgaria.

Christian M. I. M. Matthiessen and John A. Bateman. 1991. *Text generation and systemic-functional linguistics: experiences from English and Japanese*. Frances Pinter Publishers and St. Martin's Press, London and New York.

Marc Moens and Mark Steedman. 1988. Temporal ontology and temporal reference. *Computational Linguistics*, 14(2):15–28.

John Nerbonne. 1986. Reference time and time in narration. *Linguistics and Philosophy*, 9(1):83–95.

Toshiyuki Ogihara. 1995. Double-access sentences and reference to states. *Natural Language Semantics*, 3:177–210.

Rebecca Passonneau. 1988. A computational model of the semantics of tense and aspect. *Computational Linguistics*, 14(2):44–60.

Hans Reichenbach. 1947. *Elements of Symbolic Logic*. MacMillan, London.

Fei Song and Robin Cohen. 1988. The interpretation of temporal relations in narrative. In *Proceedings of the Seventh National Conference on Artificial Intelligence (AAAI-88)*, St. Paul, Minnesota.

Mark Steedman. 1995. Dynamic semantics for tense and aspect. In *The 1995 International Joint Conference on AI (IJCAI-95)*, Montreal, Quebec, Canada.

Frank Vlach. 1993. Temporal adverbials, tenses and the perfect. *Linguistics and Philosophy*, 16(3):231–283.

Bonnie Lynn Webber. 1987. The interpretation of tense in discourse. In *Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics (ACL-87)*, pages 147–154, Stanford, CA.

Guowen Yang and John Bateman. 2009. The chinese aspect generation based on aspect selection functions. In *Proceedings of the 47th Annual Meeting of the ACL and the 4th IJCNLP of the AFNLP (ACL-IJCNLP 2009)*, Singapore.