

TENSE TREES AS THE “FINE STRUCTURE” OF DISCOURSE

Chung Hee Hwang & Lenhart K. Schubert

Department of Computer Science
University of Rochester
Rochester, New York 14627, U. S. A.
{hwang,schubert}@cs.rochester.edu

ABSTRACT

We present a new compositional tense-aspect deindexing mechanism that makes use of *tense trees* as components of discourse contexts. The mechanism allows reference episodes to be correctly identified even for embedded clauses and for discourse that involves shifts in temporal perspective, and permits deindexed logical forms to be automatically computed with a small number of deindexing rules.

1 Introduction

Work on discourse structure, e.g., [Reichman, 1985; Grosz and Sidner, 1986; Allen, 1987], has so far taken a rather coarse, high-level view of discourse, mostly treating sentences or sentence-like entities (“utterance units,” “contributions,” etc.) as the lowest-level discourse elements. To the extent that sentences are analyzed at all, they are simply viewed as carriers of certain features relevant to supra-sentential discourse structure: cue words, tense, time adverbials, aspectual class, intonational cues, and others. These features are presumed to be extractable in some straightforward fashion and provide the inputs to a higher-level discourse segment analyzer.

However, sentences (or their logical forms) are not in general “flat,” with a single level of structure and features, but may contain multiple levels of clausal and adverbial embedding. This substructure can give rise to arbitrarily complex relations among the contributions made by the parts, such as temporal and discourse relations among subordinate clausal constituents and events or states of affairs they evoke. It is therefore essential, in a comprehensive analysis of discourse structure, that these intra-sentential relations be systematically brought to light and integrated with larger-scale discourse structures.

Our particular interest is in tense, aspect and other indicators of temporal structure. We are developing a uniform, compositional approach to interpretation in which a parse tree leads directly (in rule-to-rule fashion)

to a preliminary, *indexical* logical form, and this LF is *deindexed* by processing it in the current *context* (a well-defined structure). Deindexing simultaneously transforms the LF and the context: context-dependent constituents of the LF, such as operators *past*, *pres* and *perf* and adverbs like *today* or *earlier*, are replaced by explicit relations among quantified *episodes*; (anaphora are also deindexed, but this is not discussed here); and new structural components and episode tokens (and other information) are added to the context. This dual transformation is accomplished by simple recursive equivalences and equalities. The relevant context structures are called *tense trees*; these are what we propose as the “fine structure” of discourse, or at least as a key component of that fine structure.

In this paper, we first review Reichenbach’s influential work on tense and aspect. Then we describe temporal deindexing using tense trees, and extensions of the mechanism to handle discourse involving shifts in temporal perspective.

2 Farewell to Reichenbach

Researchers concerned with higher-level discourse structure, e.g., Webber [1987;1988], Passonneau [1988] and Song and Cohen [1991], have almost invariably relied on some Reichenbach [1947]-like conception of tense. The syntactic part of this conception is that there are nine tenses in English, namely *simple* past, present and future tense, past, present and future *perfect* tense, and *posterior* past, present and future tense¹ (plus progressive variants). The semantic part of the conception is that each tense specifies temporal relations among exactly three times particular to a tensed clause, namely the event time (*E*), the reference time (*R*) and the speech time (*S*). On this conception, information in discourse is a matter of “extracting” one of the nine Reichenbachian tenses from each sentence, asserting the

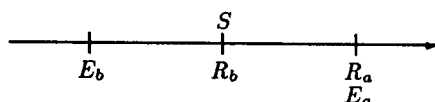
¹Examples of expressions in posterior tense are *would*, *was going to* (posterior past), *is going to* (posterior present), and *will be going to* (posterior future).

appropriate relations among E , R and S , and appropriately relating these times to previously introduced times, taking account of discourse structure cues implicit in tense shifts.

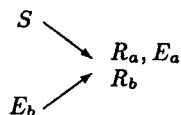
It is easy to understand the appeal of this approach when one's concern is with higher-level structure. By viewing sentences as essentially flat, carrying tense as a top-level feature with nine possible values and evoking a triplet of related times, one can get on with the higher-level processing with minimum fuss. But while there is much that is right and insightful about Reichenbach's conception, it seems to us unsatisfactory from a modern perspective. One basic problem concerns embedded clauses. Consider, for instance, the following passage.

- (1) John will find this note when he gets home.
- (2) He *will think*_(a) Mary *has left*_(b).

Reichenbach's analysis of (2) gives us $E_b < S, R_b < R_a, E_a$, where $t_1 < t_2$ means t_1 is before t_2 , as below.



That is, John will think that Mary's leaving took place some time before the speaker uttered sentence (2). This is incorrect; it is not even likely that John would know about the utterance of (2). In actuality, (2) only implies that John will think Mary's leaving took place some time before the time of his thinking, i.e., $S < R_a, E_a$ and $E_b < R_b, R_a$, as shown below.



Thus, Reichenbach's system fails to take into account the local context created by syntactic embedding.

Attempts have been made to refine Reichenbach's theory (e.g., [Hornstein, 1977; Smith, 1978; Nerbonne, 1986]), but we think the lumping together of tense and aspect, and the assignment of E, R, S triples to all clauses, are out of step with modern syntax and semantics, providing a poor basis for a systematic, compositional account of temporal relations within clauses and between clauses. In particular, we contend that English past, present, future and perfect are separate morphemes making separate contributions to syntactic structure and meaning. Note that perfect *have*, like most verbs, can occur untensed ("She is likely to have left by now"). Therefore, if the meaning of other tensed verbs such as *walks* or *became* is regarded as composite, with the tense morpheme supplying a "present" or "past" component of the meaning, the same ought to be said about tensed forms of *have*. The modals *will* and

would do not have untensed forms. Nevertheless, considerations of syntactic and semantic uniformity suggest that they too have composite meanings, present or past tense being one part and "future modality" the other. This unifies the analyses of the modals in sentences like "He knows he *will* see her again" and "He knew he *would* see her again," and makes them entirely parallel to paraphrases in terms of *going to*, viz., "He knows he is going to see her again" and "He knew he was going to see her again." We take these latter "posterior tense" forms to be patently hierarchical (e.g., *is going to see her* has 4 levels of VP structure, counting *to* as an auxiliary verb) and hence semantically composite on any compositional account. Moreover, *going to* can both subordinate, and be subordinated by, perfect *have*, as in "He is going to have left by then." This leads to additional "complex tenses" missing from Reichenbach's list.

We therefore offer a compositional account in which operators corresponding to past (*past*), present (*pres*), future (*futr*) and perfect (*perf*) contribute separately and uniformly to the meanings of their operands, i.e., formulas at the level of LF. Thus, for instance, the temporal relations implicit in "John will have left" are obtained not by extracting a "future perfect" and asserting relations among E , R and S , but rather by successively taking account of the meanings of the nested *pres*, *futr* and *perf* operators in the LF of the sentence. As it happens, each of those operators implicitly introduces exactly one *episode*, yielding a Reichenbach-like result in this case. (But note: a simple present sentence like "John is tired" would introduce only *one* episode concurrent with the speech time, not two, as in Reichenbach's analysis.) Even more importantly for present purposes, each of *pres*, *past*, *futr* and *perf* is treated uniformly in deindexing and context change. More specifically, they drive the generation and traversal of *tense trees* in deindexing.

3 Tense Trees

Tense trees provide that part of a discourse context structure² which is needed to interpret (and deindex) temporal operators and modifiers within the logical form of English sentences. They differ from simple lists of Reichenbachian indices in that they organize episode tokens (for described episodes and the utterances themselves) in a way that *echoes the hierarchy of temporal and modal operators* of the sentences and clauses from which the tokens arose. In this respect, they are anal-

²In general, the context structure would also contain speaker and hearer parameters, temporal and spatial frames, and tokens for salient referents other than episodes, among other components—see [Allen, 1987].

ogous to larger-scale representations of discourse structure which encode the hierarchic segment structure of discourse. (As will be seen, the analogy goes further.) Tense trees for successive sentences are “overlaid” in such a way that related episode tokens typically end up as adjacent elements of lists at tree nodes. The traversal of trees and the addition of new tokens is simply and fully determined by the logical forms of the sentences being interpreted.

The major advantage of tense trees is that they allow simple, systematic interpretation (by deindexing) of tense, aspect, and time adverbials in texts consisting of arbitrarily complex sentences, and involving implicit temporal reference across clause and sentence boundaries. This includes certain relations implicit in the ordering of clauses and sentences. As has been frequently observed, for a sequence of sentences within the same discourse segment, the temporal reference of a sentence is almost invariably connected to that of the previous sentence in some fashion. Typically, the relation is one of temporal precedence or concurrency, depending on the *aspectual class* or *aktionsart* involved (cf., “John closed his suitcase; He walked to the door” *versus* “John opened the door; Mary was sleeping”). However, in “Mary got in her Ferrari. She bought it with her own money,” the usual temporal precedence is reversed (based on world knowledge). Also, other discourse relations could be implied, such as *cause-of*, *explains*, *elaborates*, etc. (more on this later). Whatever the relation may be, finding the right pair of episodes involved in such relations is of crucial importance for discourse understanding. Echoing Leech [1987, p41], we use the predicate constant *orients*, which subsumes all such relations. Note that the *orients* predications can later be used to make probabilistic or default inferences about the temporal or causal relations between the two episodes, based on their aspectual class and other information. In this way they supplement the information provided by larger-scale discourse segment structures. We now describe tense trees more precisely.

Tense Tree Structure

The form of a tense tree is illustrated in Figure 1. As an aid to intuition, the nodes in Figure 1 are annotated with simple sentences whose indexical LFs would lead to those nodes in the course of deindexing. A tense tree node may have up to three branches—a leftward *past* branch, a downward *perfect* branch, and a rightward *future* branch. Each node contains a stack-like list of recently introduced episode tokens (which we will often refer to simply as episodes).

In addition to the three branches, the tree may have (horizontal) *embedding links* to the roots of embedded tense trees. There are two kinds of these embedding links, both illustrated in Figure 1. One kind,

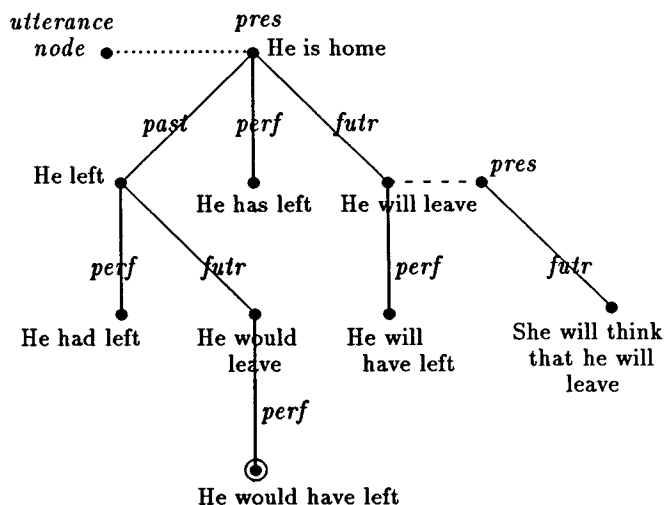


Figure 1. A Tense Tree

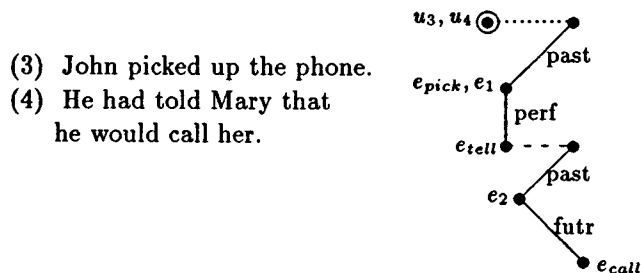
indicated by dashed lines, is created by subordinating constructions such as VPs with *that*-complement clauses. The other kind, indicated by dotted lines, is derived from the surface speech act (e.g., telling, asking or requesting) implicit in the mood of a sentence. On our view, the utterances of a speaker (or sentences of a text, etc.) are ultimately to be represented in terms of modal predications expressing these surface speech acts, such as [Speaker tell Hearer (That Φ)] or [Speaker ask Hearer (Whether Φ)]. Although these speech acts are not explicitly part of what the *speaker* uttered, they are part of what the *hearer* gathers from an utterance. *Speaker* and *Hearer* are indexical constants to be replaced by the speaker(s) and the hearer(s) of the utterance context. The two kinds of embedding links require slightly different tree traversal techniques as will be seen later.

A set of trees connected by embedding links is called a *tense tree structure* (though we often refer loosely to tense tree structures as tense trees). This is in effect a tree of tense trees, since a tense tree can be embedded by only one other tree. At any time, exactly one node of the tense tree structure for a discourse is in focus, and the focal node is indicated by $\odot$. Note that the “tense tree” in Figure 1 is in fact a tense tree structure, with the lowest node in focus.

By default, an episode added to the right end of a list at a node is “oriented” by the episode which was previously rightmost. For episodes stored at different nodes, we can read off their temporal relations from the tree roughly as follows. At any given moment, for a pair of episodes e and e' that are rightmost at nodes n and n' , respectively, where n' is a daughter of n , if the branch connecting the two nodes is a past branch, [e'

before e]³; if it is a perfect branch, [e' impinges-on e] (as we explain later, this yields entailments [e' before e] if e' is nonstative and [e' until e] if e' is stative, respectively illustrated by “John has left” and “John has been working”); if it is a future branch, [e' after e]; and if it is an embedding link, [e' at-about e]. These orienting relations and temporal relations are not extracted *post hoc*, but rather are automatically asserted in the course of deindexing using the rules shown later.

As a preliminary example, consider the following passage and a tense tree annotated with episodes derived from it by our deindexing rules:



u_3 and u_4 are utterance episodes for sentences (3) and (4) respectively.

Intuitively, the temporal content of sentence (4) is that the event of John's *telling*, e_{tell} , took place *before* some time e_1 , which is at the same time as the event of John's *picking* up the phone, e_{pick} ; and the event of John's *calling*, e_{call} , is located *after* some time e_2 , which is the at the same time as the event of John's *telling*, e_{tell} . For the most part, this information can be read off directly from the tree: [e_{pick} orients e_1], [e_{tell} before e_1] and [e_{call} after e_2]. In addition, the deindexing rules yield [e_2 same-time e_{tell}]. From this, one may infer [e_{tell} before e_{pick}] and [e_{call} after e_{tell}], assuming that the *orients* relation defaults to *same-time* here.

How does [e_{pick} orients e_1] default to [e_{pick} same-time e_1]? In the tense tree, e_1 is an episode evoked by the past tense operator which is part of the meaning of *had* in (4). It is a *stative* episode, since this past operator logically operates on a sentence of form (perf Φ), and such a sentence describes a *state* in which Φ has occurred—in this instance, a state in which John has told Mary that he will call her. It is this stativity of e_1 which (by default) leads to a *same-time* interpretation of *orients*.⁴ Thus, on our account, the tendency of past perfect “reference time” to align itself with a

previously introduced past event is just an instance of a general tendency of stative episodes to align themselves with their orienting episode. This is the same tendency noted previously for “John opened the door. Mary was sleeping.” We leave further comments about particularizing the *orients* relation to a later subsection.

We remarked that the relation [e_2 same-time e_{tell}] is obtained directly from the deindexing rules. We leave it to the reader to verify this in detail (see **Past** and **Futr** rules stated below). We note only that e_2 is evoked by the past tense component of *would* in (4), and denotes a (possible) *state* in which John will call Mary. Its stativity, and the fact that the subordinate clause in (4) is “past-dominated,”⁵ causes [e_2 bef_T e_{tell}] to be deindexed to [e_2 same-time e_{tell}].

We now show how tense trees are modified as discourse is processed, in particular, how episode tokens are stored at appropriate nodes of the tense tree, and how deindexed LFs, with *orients* and temporal ordering relations incorporated into them, are obtained.

Processing of Utterances

The processing of the (indexical) LF of a new utterance always begins with the root node of the current tense tree (structure) in focus. The processing of the top-level operator immediately pushes a token for the surface speech act onto the episode list of the root node.

Here is a typical indexical LF:

(decl (past [John know (That
(past (¬ (perf [Mary leave])))])])])
“John knew that Mary had not left.”

(*decl* stands for *declarative*; its deindexing rule introduces the surface speech act of type “tell”). As mentioned earlier, our deindexing mechanism is a compositional one in which operators *past*, *futr*, *perf*, $\neg$, *That*, *decl*, etc., contribute separately to the meaning of their operands. As the LF is recursively transformed, the tense and aspect operators encountered, *past*, *perf* and *futr*, in particular, cause the focus to shift “downward” along existing branches (or new ones if necessary). That is, processing a *past* operator shifts the current focus down to the left, creating a new branch if necessary. The resulting tense tree is symbolized as $\swarrow$ T. Similarly *perf* shifts straight down, and *futr* shifts down to the right, with respective results $\downarrow$ T and $\searrow$ T. *pres* maintains the current focus. Certain operators embed new trees at the current node, written $\mapsto$ T (e.g., *That*), or shift focus to an existing embedded tree, written $\hookrightarrow$ T (e.g., *decl*). Focus shifts to a parent or embedding node are symbolized as $\uparrow$ T and $\leftarrow$ T respectively. As a final tree operation, OT denotes storage of episode token e_T (a new episode symbol not yet used in T) at the current

³Or, sometimes, *same-time* (cf., “John noticed that Mary looked pale” vs. “Mary realized that someone broke her vase”). This is not decided in an *ad hoc* manner, but as a result of systematically interpreting the context-charged relation *bef_T*. More on this later.

⁴More accurately, the default interpretation is [(end-of e_{pick}) same-time e_1], in view of examples involving a longer preceding event, such as “John painted a picture. He was pleased with the result.”

⁵A node is *past-dominated* if there is a *past* branch in its ancestry (where embedding links also count as ancestry links).

focus, as rightmost element of its episode list. As each node comes into focus, its episode list and the lists at certain nodes on the same tree path provide explicit reference episodes in terms of which *past*, *pres*, *futr*, *perf*, time adverbials, and implicit “orienting” relations are rewritten nonindexically. Eventually the focus returns to the root, and at this point, we have a nonindexical LF, as well as a modified tense tree.

Deindexing Rules

Before we proceed with an example, we show some of the basic deindexing rules here.⁶ In the following, “**” is an episodic operator that connects a formula with the situation it characterizes. Predicates are infixed and quantifiers have restrictions (following a colon).⁷

- Decl:** $(\text{decl } \Phi)_T$
 $\leftrightarrow (\exists e_T: [e_T \text{ same-time } \text{Now}_T] \wedge$
 $\quad [\text{Last}_T \text{ immediately-precedes } e_T] \wedge$
 $\quad [\text{Speaker tell Hearer (That } \Phi_{\text{OT}})]$
 $\quad \text{** } e_T])$
Tree transform: $(\text{decl } \Phi) \cdot T = \leftarrow (\Phi \cdot (\hookrightarrow \text{OT}))$
- Pres:** $(\text{pres } \Phi)_T$
 $\leftrightarrow (\exists e_T: [e_T \text{ at-about } \text{Emb}_T] \wedge [\text{Last}_T \text{ orients } e_T])$
 $\quad [\Phi_{\text{OT}} \text{ ** } e_T])$
Tree transform: $(\text{pres } \Phi) \cdot T = (\Phi \cdot (\text{OT}))$
- Past:** $(\text{past } \Phi)_T$
 $\leftrightarrow (\exists e_T: [e_T \text{ bef}_T \text{ Emb}_T] \wedge [\text{Last}_T \text{ orients } e_T])$
 $\quad [\Phi_{\text{OT}} \text{ ** } e_T])$
Tree transform: $(\text{past } \Phi) \cdot T = \uparrow (\Phi \cdot (\text{OT}))$
- Futr:** $(\text{futr } \Phi)_T$
 $\leftrightarrow (\exists e_T: [e_T \text{ after } \text{Emb}_T] \wedge [\text{Last}_T \text{ orients } e_T])$
 $\quad [\Phi_{\text{OT}} \text{ ** } e_T])$
Tree transform: $(\text{futr } \Phi) \cdot T = \uparrow (\Phi \cdot (\text{OT}))$
- Perf:** $(\text{perf } \Phi)_T$
 $\leftrightarrow (\exists e_T: [e_T \text{ impinges-on } \text{Last}_T] \wedge$
 $\quad [\text{Last}_T \text{ orients } e_T])$
 $\quad [\Phi_{\text{OT}} \text{ ** } e_T])$
Tree transform: $(\text{perf } \Phi) \cdot T = \uparrow (\Phi \cdot (\text{OT}))$
- That:** $(\text{That } \Phi)_T \leftrightarrow (\text{That } \Phi_{\text{OT}})$
Tree transform: $(\text{That } \Phi) \cdot T = \leftarrow (\Phi \cdot (\text{OT}))$

As mentioned earlier, *Speaker* and *Hearer* in the Decl-rule are to be replaced by the speaker(s) and the hearer(s) of the utterance. Note that each equivalence pushes the dependence on context one level deeper into the LF, thus deindexing the top-level operator. The

symbols Now_T , Last_T and Emb_T refer respectively to the speech time for the most recent utterance in T , the last-stored episode at the current focal node, and the last-stored episode at the current embedding node. *bef_T* in the Past-rule will be replaced by either *before* or *same-time*, depending on the aspectual class of its first argument and on whether the focal node of T is past-dominated. In the Perf-rule, Last_T is analogous to the Reichenbachian reference time for the perfect. The *impinges-on* relation confines its first argument e_T (the situation or event described by the sentential operand of *perf*) to the temporal region preceding the second argument. As in the case of *orients*, its more specific import depends on the aspectual types of its arguments. If e_T is a stative episode, *impinges-on* entails that the state or process involved persists to the reference time (episode), i.e., $[e_T \text{ until } \text{Last}_T]$. If e_T is an event (e.g., an accomplishment), *impinges-on* entails that it occurred some-time before the reference time, i.e., $[e_T \text{ before } \text{Last}_T]$, and (by default) its main effects persist to the reference time.⁸

An Example

To see the deindexing mechanism at work, consider now sentences (5a) and (6a).

- (5) a. John went to the hospital.
 b. $(\text{decl } \uparrow_a (\text{past } \uparrow_b [\text{John goto Hospital}])) \uparrow_c$
 c. $(\exists e_1: [e_1 \text{ same-time } \text{Now}_1] \wedge$
 $\quad [[\text{Speaker tell Hearer (That$
 $\quad (\exists e_2: [e_2 \text{ before } e_1]$
 $\quad \quad [[\text{John goto Hospital}]] \text{ ** } e_2]])]$
 $\quad \text{** } e_1])$
- (6) a. The doctor told John he had broken his ankle.
 b. $(\text{decl } \uparrow_d (\text{past } \uparrow_e [\text{Doctor tell John (That$
 $\quad (\text{past } \uparrow_g (\text{perf } \uparrow_h [\text{John break Ankle}])))]) \uparrow_i$
 c. $(\exists e_3: [e_3 \text{ same-time } \text{Now}_2] \wedge$
 $\quad [e_1 \text{ immediately-precedes } e_3])$
 $\quad [[\text{Speaker tell Hearer (That$
 $\quad (\exists e_4: [e_4 \text{ before } e_3] \wedge [e_2 \text{ orients } e_4])$
 $\quad \quad [[\text{Doctor tell John (That$
 $\quad \quad (\exists e_5: [e_5 \text{ same-time } e_4]$
 $\quad \quad \quad [(\exists e_6: [e_6 \text{ before } e_5]$
 $\quad \quad \quad \quad [[\text{John break Ankle}]] \text{ ** } e_6)]$
 $\quad \quad \quad \quad \text{** } e_5)])]$
 $\quad \quad \text{** } e_4)])]$
 $\quad \text{** } e_3])$

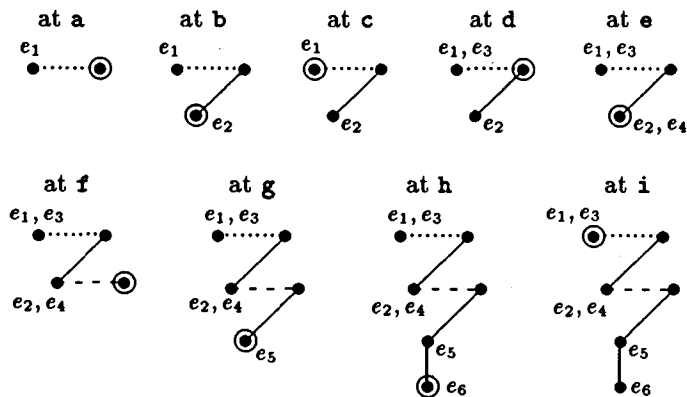
⁶ See [Hwang, 1992] for the rest of our deindexing rules. Some of the omitted ones are: *Fpres* (“futural present,” as in “John has a meeting tomorrow”), *Prog* (progressive aspect), *Pred* (predication), *K*, *Ka* and *Ke* (“kinds”), those for deindexing various operators (especially, negation and adverbials), etc.

⁷ For details of *Episodic Logic*, our semantic representation, see [Schubert and Hwang, 1989; Hwang and Schubert, 1991].

⁸ We have formulated tentative meaning postulates to this effect but cannot dwell on the issue here. Also, we are setting aside certain well-known problems involving temporal adverbials in perfect sentences, such as the inadmissibility of “*John has left yesterday.” For a possible approach, see [Schubert and Hwang, 1990].

The LFs before deindexing are shown in (5,6b) (where the labelled arrows mark points we will refer to); the final, context-independent LFs are in (5,6c). The transformation from (b)'s to (c)'s and the corresponding tense tree transformations are done with the deindexing rules shown earlier. Anaphoric processing is presupposed here.

The snapshots of the tense tree while processing (5b) and (6b), at points $\uparrow_a$ – $\uparrow_i$, are as follows (with a null initial context).



The resultant tree happens to be unary, but additional branches would be added by further text, e.g., a future branch by “It will take several weeks to heal.”

What is important here is, first, that Reichenbach-like relations are introduced compositionally; e.g., [e_6 before e_5], i.e., the breaking of the ankle, e_6 , is *before* the state John is in at the time of the doctor’s talking to him, e_4 . In addition, the recursive rules take correct account of embedding. For instance, the embedded present perfect in a sentence such as “John will think that Mary has left” will be correctly interpreted as relativized to John’s (future) thinking time, rather than the speech time, as in a Reichenbachian analysis.

But beyond that, episodes evoked by successive sentences, or by embedded clauses within the same sentence, are correctly connected to each other. In particular, note that the orienting relation between John’s going to the hospital, e_2 , and the doctor’s diagnosis, e_4 , is automatically incorporated into the deindexed formula (6c). We can plausibly *particularize* this orienting relation to [e_4 after e_2], based on the aspectual class of “goto” and “tell” (see below). Thus we have established inter-clausal connections automatically, which in other approaches require heuristic discourse processing. This was a primary motivation for tense trees. Our scheme is easy to implement, and has been successfully used in the TRAINS interactive planning advisor at Rochester [Allen and Schubert, 1991].

More on Particularizing the ORIENTS Relation

The *orients* relation is essentially an indicator that

there could be a more specific discourse relation between the argument episodes. As mentioned, it can usually be particularized to one or more temporal, causal, or other “standard” discourse relation. Existing proposals for getting these discourse relations right appear to be of two kinds. The first uses the aspectual classes of the predicates involved to decide on discourse relations, especially temporal ones, e.g., [Partee, 1984], [Dowty, 1986] and [Hinrichs, 1986]. The second approach emphasizes inference based on world knowledge, e.g., [Hobbs, 1985] and [Lascarides and Asher, 1991; Lascarides and Oberlander, 1992]. The work by Lascarides *et al.* is particularly interesting in that it makes use of a default logic and is capable of retracting previously inferred discourse relations.

Our approach fully combines the use of aspectual class information and world knowledge. For example, in “Mary got in her Ferrari. She bought it with her own money,” the successively reported “achievements” are by default in chronological order. Here, however, this default interpretation of *orients* is reversed by world knowledge: one owns things *after* buying them, rather than before. But sometimes world knowledge is mute on the connection. For instance, in “John raised his arm. A great gust of wind shook the trees,” there seems to be no world knowledge supporting temporal adjacency or a causal connection. Yet we tend to infer both, perhaps attributing magical powers to John (precisely because of the lack of support for a causal connection by world knowledge). So in this case default conclusions based on *orients* seem decisive. In particular, we would assume that if e and e' are nonstative episodes,⁹ where e is the performance of a volitional action and e' is not, then [e orients e'] suggests [e right-before e'] and (less firmly) [e cause-of e'].¹⁰

4 Beyond Sentence Pairs

The tense tree mechanism, and particularly the way in which it automatically supplies orienting relations, is well suited for longer narratives, including ones with tense shifts. Consider, for example, the following (slightly simplified) text from [Allen, 1987, p400]:

- (7) a. Jack and Sue went_{e₁} to a hardware store
- b. as someone had_{e₂} stolen_{e₃} their lawnmower.
- c. Sue had_{e₄} seen_{e₅} a man take it

⁹Non-statives could be achievements, accomplishments, culminations, etc. Our aspectual class system is not entirely settled yet, but we expect to have one similar to that of [Moens and Steedman, 1988].

¹⁰Our approach to plausible inference in episodic logic in general, and to such default inferences in particular, is probabilistic (see [Schubert and Hwang, 1989; Hwang, 1992]). The hope is that we will be able to “weigh the evidence” for or against alternative discourse relations (as particularizations of *orients*).

- d. and had_{e₆} chased_{e₇} him down the street,
- e. but he had_{e₈} driven_{e₉} away in a truck.
- f. After looking_{e₁₀} in the store, they realized_{e₁₁} that they couldn't afford_{e₁₂} a new one.

Even though {b–e} would normally be considered a subsegment of the main discourse {a, f}, both the temporal relations *within* each segment and the relations *between* segments (i.e., that the substory temporally precedes the main one) are automatically captured by our rules. For instance, e_1 and e_{11} are recognized as successive episodes, both preceded at some time in the past by e_3, e_5, e_7 , and e_9 , in that order.

This is not to say that our tense tree mechanism obviates the need for larger-scale discourse structures. As has been pointed out by Webber [1987;1988] and others, many subnarratives introduced by a past perfect sentence may continue in simple past. The following is one of Webber's examples:

- (8) a. I was_{e₁} at Mary's house yesterday.
- b. We talked_{e₂} about her sister Jane.
- c. She had_{e₃} spent_{e₄} five weeks in Alaska with two friends.
- d. Together, they climbed_{e₅} Mt. McKinley.
- e. Mary asked_{e₆} whether I would want to go to Alaska some time.

Note the shift to simple past in *d*, though as Webber points out, past perfect *could* have been used. The abandonment of the past perfect in favor of simple past signals the temporary abandonment of a perspective anchored in the main narrative – thus bringing readers “closer” to the scene (a *zoom-in* effect). In such cases, the tense tree mechanism, unaided by a notion of higher-level discourse segment structure, would derive incorrect temporal relations such as [e_5 orients e_6] or [e_6 right-after e_5].

We now show possible deindexing rules for perspective shifts, assuming for now that such shifts are independently identifiable, so that they can be incorporated into the indexical LFs. *new-pers* is a sentence operator initiating a perspective shift for its operand, and *prev-pers* is a sentence (with otherwise no content) which gets back to the previous perspective. Recent_T is the episode most recently stored in the subtree immediately embedded by the focal node of T .

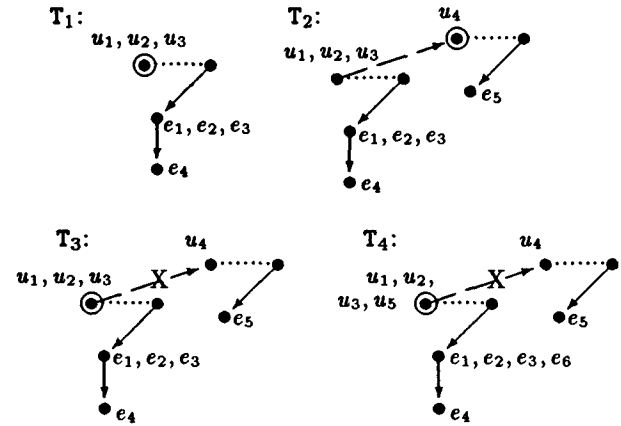
New-pers: $(\text{new-pers } \Phi)_T$
 $\leftrightarrow [\Phi \vdash T \wedge [\text{Recent}_T \text{ orients } \text{Recent}_{T'}]]$
 where $T' = \Phi \cdot (\mapsto T)$
Tree transform:
 $(\text{new-pers } \Phi) \cdot T = \Phi \cdot (\mapsto T)$

Prev-pers: $\text{prev-pers}_T \rightarrow T$ (True)
Tree transform: $\text{prev-pers} \cdot T = \leftrightarrow T$

When *new-pers* is encountered, a new tree is created and embedded at the focal node, the focus is moved to the root node of the new tree, and the next sentence is processed in that context. In contrast with other operators, *new-pers* causes an overall focus shift to the new tree, rather than returning the focus to the original root. Note that the predication [Recent_T orients $\text{Recent}_{T'}$] connects an episode of the new sentence with an episode of the previous sentence. *prev-pers* produces a trivial *True*, but it returns the focus to the embedding tree, simultaneously blocking the link between the embedding and the embedded tree (as emphasized by use of $\leftrightarrow$ instead of $\mapsto$).

We now illustrate how tense trees get modified over perspective changes, using (8) as example. We repeat (8d,e) below, augmenting them with perspective changes, and show snapshots of the tense trees at the points marked. In the trees, $u_1, \dots, u_5$ are utterance episodes for sentences a, ..., e, respectively.

- (8) d. $\uparrow_{T_1}(\text{new-pers } \text{Together, they climbed}_{\{e_5\}} \text{ Mt. McKinley.}) \uparrow_{T_2}$
 $\text{prev-pers } \uparrow_{T_3}$
- e. Mary asked_{e₆} whether I would want to go to Alaska some time. $\uparrow_{T_4}$



Notice the blocked links to the embedded tree in T_3 and T_4 . Also, note that $\text{Recent}_{T_1} = e_4$ and $\text{Recent}_{T_2} = e_5$. So, by *New-pers*, we get [e_4 orients e_5], which can be later particularized to [e_5 during e_4]. It is fairly obvious that the placement of *new-pers* and *prev-pers* operators is fully determined by discourse segment boundaries (though not in general coinciding with them). So, as long as the higher-level discourse segment structure is known, our perspective rules are easily applied. In that sense, the higher-level structure supplements the “fine structure” in a crucial way.

However, this leaves us with a serious problem: deindexing and the context change it induces is supposed to be *independent* of “plausible inferencing”; in fact,

it is intended to *set the stage* for the latter. Yet the determination of higher-level discourse structure—and hence of perspective shifts—is unquestionably a matter of plausible inference. For example, if *past perfect* is followed by *past*, this could signal either a new perspective within the current segment (see 8c,d), or the closing of the current subsegment with no perspective shift (see 7e,f). If *past* is followed by *past*, we may have either a continuation of the current perspective and segment (see 9a,b below), or a perspective shift with opening of a new segment (see 9b,c), or closing of the current segment, with resumption of the previous perspective (see 9c,d).

- (9) a. Mary found that her favorite vase was broken.
- b. She was upset.
- c. She bought it at a special antique auction,
- d. and she was afraid she wouldn't be able to find anything that beautiful again.

Only plausible inference can resolve these ambiguities. This inference process will interact with resolution of anaphora and introduction of new individuals, identification of spatial and temporal frames, the presence of modal/cognition/perception verbs, and most of all will depend on world knowledge. In (9), for instance, one may have to rely on the knowledge that one normally would not buy broken things, or that one does not buy things one already owns.

As approaches to this general difficulty, we are thinking of the following two strategies: (A) Make a best initial guess about presence or absence of *new-pers/prev-pres*, based on *surface* (syntactic) cues and then use failure-driven backtracking if the resulting interpretation is incoherent. A serious disadvantage would be lack of integration with other forms of disambiguation. (B) Change the interpretation of *Last_T*, in effect providing multiple alternative referents for the first argument of *orients*. In particular, we might use

Last_T = { e_i | e_i is the last-stored episode at the focus of *T*, or was stored in the subtree rooted at the focus of *T* after the last-stored episode at the focus of *T*}.

Subsequent processing would resemble anaphora disambiguation. In the course of further interpreting the deindexed LF, plausible inference would particularize the schematic orienting relation to a temporal (or causal, etc.) relation involving just two episodes. The result would then be used to make certain structural changes to the tense tree (*after* LF deindexing).

For instance, suppose such a schematic orienting relation is computed for a simple past sentence following a past perfect sentence (like 8c,d). Suppose further that the most coherent interpretation of the second sentence

(i.e., 8d) is one that disambiguates the orienting relation as a simple temporal inclusion relation between the successively reported events. One might then move the event token for the second event (reported in simple past) from its position at the past node to the rightmost position at the past perfect node, just as if the second event had been reported in the past perfect. (One might in addition record a perspective shift, if this is still considered useful.) In other words, we would “repair” the distortion of the tense tree brought about by the speaker’s “lazy” use of simple past in place of past perfect. Then we would continue as before.

In both strategies we have assumed a general coherence-seeking plausible inference process. While it is clear that the attainment of coherence entails delineation of discourse segment structure and of all relevant temporal relations, it remains unclear in which direction the information flows. Are there independent principles of discourse and temporal structure operating *above* the level of syntax and LF, *guiding* the achievement of full understanding, or are higher-level discourse and temporal relations a mere byproduct of full understanding? Webber [1987] has proposed independent temporal focusing principles similar to those in [Grosz and Sidner, 1986] for discourse. These are not deterministic, and Song and Cohen [1991] sought to add heuristic constraints as a step toward determinism. For instance, one constraint is based on the presumed incoherence of simple present followed by past perfect or posterior past. But there are counterexamples; e.g., “Mary is angry about the accident. The other driver had been drinking.” Thus, we take the question about independent structural principles above the level of syntax and LF to be still open.

5 Conclusion

We have shown that tense and aspect can be analyzed compositionally in a way that accounts not only for their more obvious effects on sentence meaning but also, via tense trees, for their cumulative effect on context and the temporal relations implicit in such contexts. As such, the analysis seems to fit well with higher-level analyses of discourse segment structure, though questions remain about the flow of information between levels.

Acknowledgements

We gratefully acknowledge helpful comments by James Allen and Philip Harrison on an earlier draft and much useful feedback from the members of TRAINS group at the University of Rochester. This work was supported in part by NSERC Operating Grant A8818 and

ONR/DARPA research contract no. N00014-82-K-0193, and the Boeing Co. under Purchase Contract W-288104.

A preliminary version of this paper was presented at the AAAI Fall Symposium on Discourse Structure in Natural Language Understanding and Generation, Pacific Grove, CA, November 1991.

References

- [Allen, 1987] J. Allen, *Natural Language Understanding*, Chapter 14. Benjamin/Cummings Publ. Co., Reading, MA.
- [Allen and Schubert, 1991] J. Allen and L. K. Schubert, "The TRAINS project," TR 382, Dept. of Comp. Sci., U. of Rochester, Rochester, NY.
- [Dowty, 1986] D. Dowty, "The effect of aspectual classes on the temporal structure of discourse: semantics or pragmatics?" *Linguistics and Philosophy*, 9(1):37-61.
- [Grosz and Sidner, 1986] B. J. Grosz and C. L. Sidner, "Attention, intentions, and the structure of discourse," *Computational Linguistics*, 12:175-204.
- [Hinrichs, 1986] E. Hinrichs, "Temporal anaphora in discourses of English," *Linguistics and Philosophy*, 9(1):63-82.
- [Hobbs, 1985] J. R. Hobbs, "On the coherence and structure of discourse," Technical Report CSLI-85-37, Stanford, CA.
- [Hornstein, 1977] N. Hornstein, "Towards a theory of tense," *Linguistic Inquiry*, 3:521-557.
- [Hwang, 1992] C. H. Hwang, *A Logical Framework for Narrative Understanding*, PhD thesis, U. of Alberta, Edmonton, Canada, 1992, To appear.
- [Hwang and Schubert, 1991] C. H. Hwang and L. K. Schubert, "Episodic Logic: A situational logic for natural language processing," In *3rd Conf. on Situation Theory and its Applications (STA-3)*, Oiso, Kanagawa, Japan, November 18-21, 1991.
- [Lascarides and Asher, 1991] A. Lascarides and N. Asher, "Discourse relations and defeasible knowledge," In *Proc. 29th Annual Meeting of the ACL*, pages 55-62. Berkeley, CA, June 18-21, 1991.
- [Lascarides and Oberlander, 1992] A. Lascarides and J. Oberlander, "Temporal coherence and defeasible knowledge," *Theoretical Linguistics*, 8, 1992, To appear.
- [Leech, 1987] G. Leech, *Meaning and the English Verb (2nd ed)*, Longman, London, UK.
- [Moens and Steedman, 1988] M. Moens and M. Steedman, "Temporal ontology and temporal reference," *Computational Linguistics*, 14(2):15-28.
- [Nerbonne, 1986] J. Nerbonne, "Reference time and time in narration," *Linguistics and Philosophy*, 9(1):83-95.
- [Partee, 1984] B. Partee, "Nominal and Temporal Anaphora," *Linguistics and Philosophy*, 7:243-286.
- [Passonneau, 1988] R. J. Passonneau, "A Computational model of the semantics of tense and aspect," *Computational Linguistics*, 14(2):44-60.
- [Reichenbach, 1947] H. Reichenbach, *Elements of Symbolic Logic*, Macmillan, New York, NY.
- [Reichman, 1985] R. Reichman, *Getting Computers to Talk Like You and Me*, MIT Press, Cambridge, MA.
- [Schubert and Hwang, 1989] L. K. Schubert and C. H. Hwang, "An Episodic knowledge representation for Narrative Texts," In *Proc. 1st Inter. Conf. on Principles of Knowledge Representation and Reasoning (KR '89)*, pages 444-458, Toronto, Canada, May 15-18, 1989. Revised, extended version available as TR 345, Dept. of Comp. Sci., U. of Rochester, Rochester, NY, May 1990.
- [Schubert and Hwang, 1990] L. K. Schubert and C. H. Hwang, "Picking reference events from tense trees: A formal, implementable theory of English tense-aspect semantics," In *Proc. Speech and Natural Language, DARPA Workshop*, pages 34-41, Hidden Valley, PA, June 24-27, 1990.
- [Smith, 1978] C. Smith, "The syntax and interpretations of temporal expressions in English," *Linguistics and Philosophy*, 2:43-99.
- [Song and Cohen, 1991] F. Song and R. Cohen, "Tense interpretation in the context of narrative," In *Proc. AAAI-91*, pages 131-136. Anaheim, CA, July 14-19, 1991.
- [Webber, 1987] B. L. Webber, "The Interpretation of tense in discourse," In *Proc. 25th Annual Meeting of the ACL*, pages 147-154, Stanford, CA, July 6-9, 1987.
- [Webber, 1988] B. L. Webber, "Tense as discourse anaphor," *Computational Linguistics*, 14(2):61-73.