

Picking reference events from tense trees: a formal, implementable theory of English tense–aspect semantics

Article · January 1990

DOI: 10.3115/116580.116597

2 authors:



Lenhart K. Schubert
University of Rochester



Chung Hee Hwang
General Electric

Picking Reference Events from Tense Trees: A Formal, Implementable Theory of English Tense-Aspect Semantics

Lenhart K. Schubert and Chung Hee Hwang

Department of Computer Science, University of Rochester
Rochester, New York 14627, U. S. A., and
Department of Computing Science, University of Alberta
Edmonton, Alberta, Canada T6G 2H1

Abstract. Despite numerous investigations into English tense-aspect semantics, the problem of formally interpreting tense and aspect remains in large part unsolved. A formal solution requires (a) a well-defined mapping from a subset of English covering the most common tense-aspect constructions (including embedded tensed clauses) to a formal meaning representation, and (b) a well-defined denotational semantics for the meaning representation, which accords with speakers' intuitions about the original text. We propose a simple structure called a *tense tree* (or a set of connected tense trees) as a basis for interpreting tense-aspect constructions and some time adverbials in a context-dependent way. The de-indexicalization process simultaneously transforms tense trees and logical forms, the former in accord with simple recursion equations and the latter in accord with formal equivalences between context-indexed and context-independent logical forms. The rules have been implemented, and yield meaning representations in a formal episodic logic for narrative understanding.

I. Introduction

Narratives describe and relate *episodes* (events, situations, eventualities). The episodes successively introduced appear to “line up” in systematic ways as a function of the text structure, with tense and aspect playing a crucial role. (1-2) illustrate this familiar phenomenon:

- (1) John grinned at Mary. She frowned.
- (2) John knocked, but soon realized Mary had left.

Each of the verbs in (1) appears to introduce an episode; the past inflection places them before the utterance event (or speech time), and the surface ordering suggests temporal sequencing, and probably a causal connection. In (2) we appear to have at least three episodes, a “knocking,” a “realizing” and a “leaving.” The past perfect auxiliary relates the “leaving” to a past episode serving as reference point. (This reference point seems closely correlated with the “realizing” episode, but on our theory is not identical with it.) As well, the adverb *soon* implicitly references and relates the “knocking” and “realizing” episodes.

The problem of automatically extracting these relationships is complicated by a number of subtle issues:

- How many episodes (events, situations, etc.) does a sentence implicitly introduce? We remarked that the introduction of episodes appears correlated with verbs; but in

- (3) John did not elude the tackle and score a touchdown, disappointing the fans.

the causally relevant episode is characterized by the *non*-occurrence of successive “eluding” and “scoring” events; and in

- (4) When each guest congratulated Mary and gave her a present, she felt quite overwhelmed.

the focus is on a “quantified” event consisting of any number of “congratulating” and “giving” events.

- What are the relative (semantic) *scopes* of tense and aspect, relative to NPs, VPs, adverbials, negation, etc.? Syntax suggests that tense scope is confined to the tensed verb or VP, yet in a sentence like

- (5) For a whole week this summer, no rain fell.

this simple view is troublesome: if the past tense is confined to the VP (so that *fell* means “falls-before-now”), then the sentence is true if we can find a week any time this summer in which there are no *past* rainfalls – a property that holds of all *remaining* (i.e., future) weeks of the summer!

- Do simple past or present-tense sentences make *definite* or *indefinite* reference to episodes (or times)? Davidson[3], Harman[5], Reichenbach[13] and a majority of writers after them took event reference in simple past and present to be indefinite (existential), and this seems in accord with intuition in such sentences as

- (6) John got married last year.

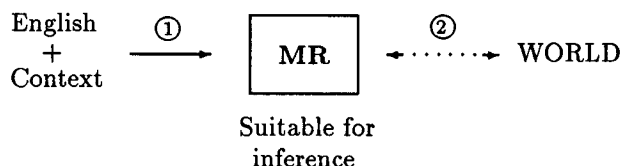
where there is no definite commitment as to the time of the event (apart from its being confined to last year). On the other hand, Partee[12] pointed out the seemingly anaphoric character of tense reference in sentences like

- (7) I left the stove on.

Also, Leech[8], McCawley[9], Webber[19] and many others have noted that the simple past normally involves a “point of orientation” such as an immedi-

ately preceding event, and that Reichenbach’s “reference time” appears to be definite at least in the case of perfect.

Our goal is a comprehensive account of how episodes and their relationships are determined by syntax, semantics, and pragmatics. The account is to be an intrinsic part of a general theory of the understanding process, and is to be formulated in a mathematically precise and computationally implementable way. At the most general level, the task can be viewed in terms of the following sort of schema:



I.e., English in conjunction with suitable context structures needs to be mapped to a formal meaning representation (①), and this MR needs to be formally interpretable (②), with truth conditions that are in intuitive accord with intuitions about the original English text. Thus we will have some assurance that inferences based on the MR will be those intuitively warranted.

With respect to this schema, most previous studies of tense and aspect have been more formally explicit about ② than about ①. The MR is often some dialect of first-order logic, or a modal or intensional logic, with formal semantics in the traditional vein (e.g., possible-worlds models with multiple time indices).

The description of ①, however, is often very informal or limited to a very restricted set of constructions. Furthermore, the formalized accounts (e.g., [4], [6], [14]) tend to take context for granted – the focus is on truth conditions, and so the model is simply assumed to supply the needed reference times. AI-motivated work on tense and aspect has emphasized their pragmatic significance, e.g., for discourse segment structure (see [1], ch. 13 & 14), and their role in the practical extraction of temporal relationships from discourse in applied NL systems (e.g., see several of the articles in *Computational Linguistics* 14(2), 1988). Much of this work has treated ① heuristically, restricting attention to simple types of sentences and viewing tense and aspect as phenomena to be handled by a separate module, rather than as an integral part of a compositional mapping from text to meaning representations.

Non-compositional approaches based on simple sentence types run the risk of being ungeneralizable. For instance, they may tie event introduction to verb phrases in a way that doesn’t generalize to logical compounds (as in sentences (3) and (4)). Also, they may get the semantics of embedded sentences wrong. For instance, simple sentence types suggest that past tense places an event before the speech time. But in

- (8) Mary will realize a year from now that her wedding to John, which takes place tomorrow, was a big mistake.

the past tense is used for a *future* event. Thus we need to characterize the semantics of particular syntactic constructs – including tense-aspect constructs – in a way that holds for *all* syntactic environments in which the constructs may occur.

On the issue of compositionality, we should also remark that we find much of the work on tense and aspect “unduly Reichenbachian”. Though the importance of Reichenbach’s seminal work in this area is undisputed, it is based on the dubious view that tensed perfects are complex variants of “simple” past, present and future tense. Since these “complex” tenses appear to involve a reference time (besides the event and speech times), Reichenbach attributed such a reference time to “simple” tenses as well. But from a compositional perspective, it is much more natural to view tense (past and present) and perfect as making *separate* contributions to the VP or sentence meaning. After all, *each* can occur in a VP without the other (“John probably *forgot*”, “John is likely to *have forgotten*”). Thus semantic dependence on Reichenbach’s reference time may well be particular to the perfect – just as dependence on speech time (or a more general sentence reference time) seems particular to tense.

In the remaining sections of the paper we report our progress so far on the tense-aspect problem within the above schema. Section 2 briefly elaborates our conception of the understanding process. Section 3 explains the intuitive idea behind our method of transforming indexical logical forms (LFs) to nonindexical LFs in a principled way, using context structures called *tense trees*. These provide the “points of orientation” (reference episodes) needed to interpret tense and aspect, and are transformed in the course of “de-indexicalizing” LFs in a way that can be concisely stated by recursion equations. Section 4 provides a formal statement with examples of the basic de-indexicalization rules for tense and aspect. Section 5 summarizes our progress and work left to be done.

In the course of the paper, we make commitments about key aspects of the issues enumerated above. Briefly, our view is that at the level of LF, tensed English sentences contain separate *past*, *pres*, *perf* and *futr* operators; (*will* and *would* have separate tense and *futr* components); that not only *past*, *pres*, *perf*, and *futr* introduce episodes, but that various logical operators do so as well, including negation, conjunction, and quantification; that *past* and *pres* are sentence-level operators generally taking wide scope (despite syntactic appearances); and that the episodes introduced by *past* and *pres*, and indeed by any operators, are *indefinite* (existentially quantified). This last claim applies *even to the past in the past perfect*, consistently with our compositional view of tense and aspect. The way the past perfect auxiliary connects up with a pre-existing “point of orientation” is not by direct anaphoric reference, but by a mechanism that applies to all statives.

II. A Formal View of the Understanding Process

The following illustrates our current view of the stages of the understanding process, at a theoretical level. (At the procedural level these stages are interleaved.) This view incorporates ideas from GPSG, HPSG, DRT and situation semantics, and evolved in the course of our own research on mapping English into logic [17], [18] and story understanding [15], [16].

0. Assume that an episode e_0 has just been described.
 e_0 : After their doubles victory, Mary and Sue were surrounded by their fans.

New sentence: A man kissed Mary

1. **GPSG parser:**
 $[s [NP [_{Det} a] [N man]] [VP [v kissed] [NP Mary]]]$
2. **Rule-by-rule LF computation:**
 $[<\exists man> <past kiss> Mary]$
3. **Scoping:** (past $(\exists x:[x man][x kiss Mary])$)
4. **De-indexicalization** (using context structure C):
 $(\exists e:[e \text{ before } Now_3] \wedge [e_0 \text{ orients } e])$
 $[(\exists x:[x man][x kiss Mary]) ** e]$
5. **Pragmatic inference:**
 $(\exists e:[e \text{ before } Now_3] \wedge [e \text{ right-after } e_0])$
 $[(\exists x:[x man][x kiss Mary]) ** e]$
6. **Inference based on MPs & world knowledge:**
 The man liked Mary a lot;
 he was happy about her victory; etc.

A preliminary unscoped, indexical LF is derived via semantic rules paired with phrase structure rules. Predicate infixing (with the predicate following its “subject” argument and followed by additional arguments, if any) is used for readability. Angle brackets indicate unscoped operators that are to be “raised” to some sentence-level position. Scoping quantifiers involves variable introduction and conversion of the restriction predicate to a restriction formula. Note that after the third stage the \exists -quantifier and *past* operator are scoped in the example; in general we take tense to have a strong, though not absolute, wide-scoping tendency (but like quantifiers, it is “trapped” by scope islands, such as embedded clauses).

The fourth stage is our main focus in this paper. In this stage the scoped, but still indexical, LF is de-indexicalized. In particular, explicit episodic variables are introduced, and tense and aspect operators are replaced by relationships among episodes. This process makes use of tense trees as part of a context structure, modifying these trees by adding branches and episode tokens as a “side effect”. In general we envisage this stage as also performing other kinds of de-indexicalization, most importantly the explicit augmentation of anaphoric expressions (such as (The $x:[x man] \dots$)) by context-derived predications (such as (The $x:[x man] \wedge [x \text{ has-possible-antecedents } (tuple \dots)] \dots$)).

Note the connective “**” in the nonindexical formulas, connecting a formula to an episode which it *characterizes* (or *completely describes*). This is a strengthened version of “*”, or *partially describes*, an operator that is essentially the same as Reichenbach[13]’s “*”, and similar to Barwise[2]’s \models . For semantics see [16].

Note also the *orients* relationship in 4th stage (italicized) sample result. This is intended to echo Leech’s notion of a “point of orientation” (as cited in [19]). We refer to this relation, and certain others derived from context (including *has-possible-antecedents* above) as *context-charged*. The idea is that their meaning is to be “discharged” using *uncertain* (probabilistic) inference. For instance, the fact that e_0 in the example *orients* (serves as point of orientation for) e *suggests* among other possibilities that e immediately follows e_0 (in e_0 ’s “consequent” or “result” phase, in the terminology of [10] or [11]). Given the assumed types of e_0 (Mary and Sue being surrounded by their fans) and e (a man kissing Mary), this is a very plausible inference, but in other cases the most plausible conclusion from the *orients* relation may be a subepisode relation (“It was a great party. Mary did her Cher impression”), an explanatory relation (“John went shopping. Mary’s birthday was in a week”), or any of the discourse relations that have been discussed in the literature (e.g., [7]).

The de-indexicalization stage is followed by inference stages which discharge context-charged relations and more generally do *input-driven* plausible inference based on the text interpreted so far, meaning postulates (MPs) and world knowledge. In these stages unique referents for referring expressions, predictions, and explanations are computed which ultimately give a causally coherent elaboration of what has been said.

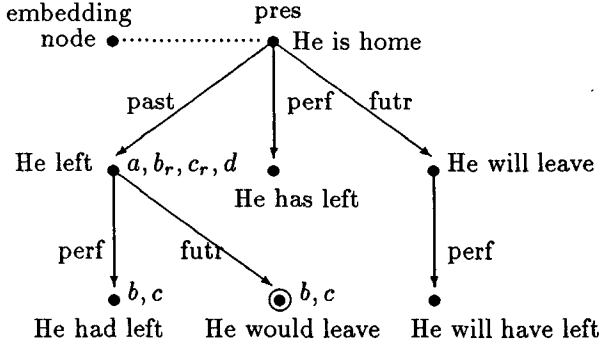
We now turn to a detailed account of the de-indexicalization stage. For a much more complete description of our overall approach to knowledge representation, inference and understanding, the reader is referred to [15] and [16].

III. De-Indexicalization: Traversing Tense Trees

A formal theory of how indexical formulas are converted to nonindexical ones as a function of context requires an explicit specification of the type of context structure to be used. We will not attempt a full specification, since we are not addressing all aspects of de-indexicalization (such as anaphoric processing).

Rather, we will specify a new type of context component called a *tense tree* (or a set of embedded trees), which serves the purpose of context-dependent tense-aspect interpretation. Apart from this, we assume a “clock” which generates a succession of *Now*-points, hearer and speaker parameters, and whatever additional information a more complete discourse theory may call for; e.g., a nested segment structure which records the evolving set of relationships among discourse segments,

Figure 1. Tense Trees



including time, hearer and speaker parameters for those segments, history lists or focus lists of entities referenced within them, etc. In fact, we may want to view our tense tree structures as part of the “fine-grained” structure of discourse segments, recording the pragmatic relationships among the clauses and subclauses of a segment.

The form of a tense tree is illustrated in Figure 1. Each node has up to three branches, a diagonally leftward one corresponding to a *past* operator, a vertical one to a *perf* operator, and a diagonally rightward one to a *futr* operator. As an indexical LF is processed in a recursive top-down manner to de-indexicalize its tense and aspect operators (and adverbials, etc.), the tense tree is traversed in a way dictated by the operators encountered. The position of the current traversal is called the *focal node*, or *focus*, of the tense tree structure and is indicated as \odot . Where branches do not exist, they are created, and as new episode variables are introduced into the LF, copies of these variables are added to lists maintained at the nodes of the tense tree. 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 de-indexicalization.

Subordinating constructions (as in VPs with *that*-complement clauses) cause one tree to *embed* another, and this is indicated by horizontal *embedding links* from a (non-root) node of a tense tree to the root node of another. A set of trees connected by embedding links is called a *tense tree structure*. (This is in effect a tree of tense trees, since a tense tree can be embedded by only one other tree.) In Figure 1 an embedding link to the root is shown, since we assume that the utterances of a speaker (or sentences of a text, etc.) are ultimately represented in terms of *performative* modal predications, such as (*pres* [*Speaker assert (that φ)*]) or (*pres* [*Speaker ask (whether φ)*])). Here *that* and *whether* are “facsimiles” of the nominalization operators *That* and *Whether*. By that we mean that they are semantically indistinguishable from the latter, but since they were in effect “inserted” by the hearer, they are treated slightly differently in the tree traversals (as will be seen).

The additional annotations a, b_r, c_r, d at the *past* node in Figure 1, and b, c at the *past-perf* and *past-futr* nodes,

are based on the following two short passages:

- (9) a. John entered the room.
 b. Mary had taken down his paintings,
 c. and had hung up Schwarzenegger posters.
 d. He groaned.
- (10) a. Mary’s note made up John’s mind.
 b. He would leave,
 c. and would let her live her own life.
 d. He began to pack.

The a at the *past* node indicates that an episode token for the event of John’s entering the room in (9) a, or the event of Mary’s note making up John’s mind in (10) a, is added at that node in de-indexicalizing the LFs of those sentences. Similarly the b_r indicates that an episode token for the reference point of the perfect in (9) b or (10) b is added at the *past* node in processing the LFs of these sentences; similarly for b, c_r, c , and d .

We will specify these operations formally, but what we have said so far allows us to explain intuitively how occurrences of tense-aspect operators are converted to explicit episode relationships. The idea is that episode tokens *adjacent* at a node (e.g., tokens for a and b_r above) generate *orienting* relationships (e.g., [a *orients* b_r]); the episode last added at a *past* node is *before* a correlated episode at the mother’s embedding node (which will be a speaking episode, thinking episode, or the like); (simplifying a bit) the episode last added at a *perf* node lasts at least *until* a correlated episode at the mother; and similarly for *futr*.

A crucial observation about (9) is that the tokens for the “entering” event in a and the “taking down paintings” event in b are *not* placed at the same node – whereas the “groaning” event in d does end up at the same node as a . Assuming that the “reference episodes” b_r and c_r can be inferred to be at approximately the same time as the entering event a (and it turns out that they can), then the “groaning” event d can be inferred to be shortly after the “entering”. As well, note that the “taking down paintings” and “hanging up Schwarzenegger posters” episodes b and c are adjacent at the *past-perf* node, so that b *orients* c . So the collocation of episode tokens in the tense tree is such that we can “read off” the relationships implicit in the tense-aspect operators and surface ordering of sentences.

We can now make the de-indexicalization process and associated tense tree transformations more precise.

IV. Basic Tense-Aspect Rules

Assume that we already have a tense tree structure T , with a particular node in focus, as a result of processing previous inputs and partially processing the current input. (If not, we generate a one-node tree using the “new-tree” function Δ .) Then de-indexicalization of an LF formula with a top-level *pres* operator, relative to T , is defined by the equivalence

$$\text{Pres: } (\text{pres } \Phi)_T \leftrightarrow (\exists e_T: [[e_T \text{ at-about Emb}_T] \wedge [\text{Last}_T \text{ orients } e_T]])$$

$$\begin{aligned} & [\Phi_{\circ\mathbf{T}} ** e_{\mathbf{T}}], \\ \text{tree transformation: } & (\text{pres } \Phi) \cdot \mathbf{T} = \Phi \cdot (\circ\mathbf{T}) \end{aligned}$$

Here $e_{\mathbf{T}}$ is assumed to be a new episode variable name uniquely defined for any given \mathbf{T} (e.g., the letter e with subscript i where i is the least integer such that e_i does not yet occur in \mathbf{T}). “ $\circ\mathbf{T}$ ” means “store $e_{\mathbf{T}}$ at the focal node of \mathbf{T} .” $\text{Emb}_{\mathbf{T}}$ denotes the last-added episode at the node which directly embeds the tree containing the focal node of \mathbf{T} . (This is usually an episode corresponding to a performative or attitude verb.) If there is no embedding node, $\text{Emb}_{\mathbf{T}}$ is the *Now*-point of the current context. $\text{Last}_{\mathbf{T}}$ is the last-stored episode variable at the focus of \mathbf{T} .

Thus the de-indexicalization rule “creates” a new episode token, which it predicates to be at about the time of embedding event (e.g., the assertion of the sentence being processed), designates the last-stored episode at the current focus as the point of orientation for the new episode, and states that the formula Φ on which *pres* operates, after recursive de-indexicalization (with $e_{\mathbf{T}}$ now stored at the focus), *characterizes* the new episode. The tree transformation which is induced by this de-indexicalization (in the implementation, as a side effect) is separately stated above. The dot symbolizes the tree-structure transformation function,

$$\begin{aligned} \cdot & : \text{LF-expressions} \times \text{Tense-tree structures} \\ & \rightarrow \text{Tense-tree structures} \end{aligned}$$

Thus, the effect of $(\text{pres } \Phi)$ on \mathbf{T} is just storage of $e_{\mathbf{T}}$ at the focus of \mathbf{T} , followed by whatever additional transformations Φ induces. (Note that the function composition in “ $\Phi \cdot (\circ\mathbf{T})$ ” is read “from the inside to the outside,” as usual.)

Next, the de-indexicalization of *past* is given by

$$\begin{aligned} \text{Past: } & (\text{past } \Phi)_{\mathbf{T}} \leftrightarrow (\exists e_{\mathbf{T}}: [[e_{\mathbf{T}} \text{ before Emb}_{\mathbf{T}}] \wedge \\ & \quad [\text{Last}_{\swarrow\mathbf{T}} \text{ orients } e_{\mathbf{T}}]] \\ & \quad [\Phi_{\circ\swarrow\mathbf{T}} ** e_{\mathbf{T}}]), \\ \text{tree transformation: } & (\text{past } \Phi) \cdot \mathbf{T} = \uparrow(\Phi \cdot (\circ\swarrow\mathbf{T})) \end{aligned}$$

Here “ $\swarrow\mathbf{T}$ ” signifies \mathbf{T} with the focus displaced to the left (i.e., *past*) daughter, with creation of a new daughter if necessary. Thus the orienting episode for the new past episode $e_{\mathbf{T}}$ is the last-stored episode at the *past* daughter of the focal node. (So for a succession of simple past-tensed sentences, each episode generated will orient the next one. As well, all of them will be before their embedding episode, i.e., the appropriate sentence utterance.) Again, the recursively de-indexicalized Φ (with the \mathbf{T} -focus shifted left and $e_{\mathbf{T}}$ stored there) is taken to *characterize* the new episode. In the tree transformation equation, the upward arrow signals upward displacement of the focus to the mother. This restores the focus to its original position, assuming that recursive processing of Φ has no net effect on the focus (which it doesn’t, thanks to the way the remaining rules work).

Actually, the above *Past*-rule is a slight simplification. It applies as stated only when the focal node is *not* “past-dominated,” i.e., if there are no *past*-branches in its ancestry (with embedding links counting as ancestry

links). If the focus is past-dominated, [$e_{\mathbf{T}}$ before $\text{Emb}_{\mathbf{T}}$] is replaced by [$e_{\mathbf{T}}$ at-or-before $\text{Emb}_{\mathbf{T}}$]. Here *at-or-before* is regarded as a context-charged relation, with different probable consequences depending on the aspectual class of its first argument. In particular, for *stative* $e_{\mathbf{T}}$, [$e_{\mathbf{T}}$ at-or-before $\text{Emb}_{\mathbf{T}}$] strongly suggests [$e_{\mathbf{T}}$ at $\text{Emb}_{\mathbf{T}}$]. This is aimed at “sequence of tense” phenomena, observable in sentences like

(11) John knew that Mary left.

(12) John knew Mary had a cold.

For (11), our rules predict that the “leaving” is strongly preferred to be prior to the “knowing” (though a variant like “John noticed that Mary winked at him” forces a concurrent-event reading); while for (12), they predict that the “cold” episode is strongly preferred to be at the same time as the “knowing” episode (though a variant like “John remembered that Mary had a cold when they got married” forces an earlier-episode reading).

To understand the next rule, recall that we take *will* and *would* (in their future and future-in-the-past readings) to consist logically of tense plus the *futr* operator. Thus the *futr* operator is encountered *only after* its implicit tense operator has been processed, so that a characterizing (“**”) relation will already embed the (*futr* Φ) expression.

$$\begin{aligned} \text{Futr: } & [(\text{futr } \Phi)_{\mathbf{T}} ** \eta] \leftrightarrow [(\exists e_{\mathbf{T}}: [[e_{\mathbf{T}} \text{ after } \eta] \wedge \\ & \quad [\text{Last}_{\searrow\mathbf{T}} \text{ orients } e_{\mathbf{T}}]] \\ & \quad [\Phi_{\circ\searrow\mathbf{T}} ** e_{\mathbf{T}}]) ** (F \eta)], \\ \text{tree transformation: } & (\text{futr } \Phi) \cdot \mathbf{T} = \uparrow(\Phi \cdot (\circ\searrow\mathbf{T})) \end{aligned}$$

This is quite analogous to *pres* and *past*, except that the temporal location of the new episode $e_{\mathbf{T}}$ is specified relative to the episode η (usually a present or past episode) characterized by “having Φ true in its future,” rather than relative to an “embedding episode.” Also, $(F \eta)$ in the rule denotes “the facts about η .” Roughly speaking, this is needed because the formula preceding “** ($F \eta$)” is an *atemporal* characterization of the propositional content of η (one that is simply true or false, rather than true of some episodes and false of others); whereas $(\text{futr } \Phi)_{\mathbf{T}}$ is a *temporal* characterization of η (“being followed at some future time by a Φ -episode” can be true of some episodes – namely, those that do precede a Φ -episode – and false of others). The “** ($F \eta$)” could be dropped without falsifying the *forward* direction of the equivalence.

For the perfect, we need to assume an ambiguity: the present perfect auxiliary (*has* or *have*) is always translated as $\lambda P \lambda x \langle \text{pres } (\text{perf}_1 [x P]) \rangle$ (so that apart from the tense, the perfect is logically perf_1); but when occurring untensed or in combination with *past*, the perfect is logically either perf_1 or perf_2 .

$$\begin{aligned} \text{Perf}(1): & [(\text{perf}_1 \Phi)_{\mathbf{T}} ** \eta] \\ & \leftrightarrow [(\exists e_{\mathbf{T}}: [e_{\mathbf{T}} \text{ until } \eta][\Phi_{\circ\mathbf{T}} @ e_{\mathbf{T}}]) ** (F \eta)], \\ \text{tree transformation: } & (\text{perf}_i \Phi) \cdot \mathbf{T} = \uparrow(\Phi \cdot (\circ\downarrow\mathbf{T})), \quad (i = 1, 2); \end{aligned}$$

where $[\Phi @ \eta] \leftrightarrow [\Phi * \eta]$, for Φ stative;

$[\Phi @ \eta] \leftrightarrow [\exists e: [e \text{ recent-subep } \eta][\Phi * e]]$, otherwise.

Here *recent-subep* is another context-charged relation, which for a *state-change* e of type Φ as first argument suggests the truth of $[\Psi * (\text{fin } \eta)]$, whenever $[\Phi ** e_1]$ entails $[\Psi * (\text{fin } e_1)]$ for all e_1 (where *fin* means “final moment of”). For instance, consider

- (13) John has been sleeping.
 (14) John has woken up.

For (13), we will get a subformula $[(\text{prog } [\text{John sleep}]) @ e_i]$, for some i . Since progressives are stative, this is equivalent to $[(\text{prog } [\text{John sleep}]) * e_i]$, meaning that John is sleeping over the entire episode e_i . This episode lasts until “ η ” in the complete formula, which is nothing else but the *reference* episode for the perfect (corresponding to Reichenbach’s reference time). For (13), this η is a “present” episode (by the *Pres*-rule), so (13) means, in effect, that John has been sleeping *until now*. We take this to be the desired inference.

For (14), we will get a subformula $[[\text{John wake-up}] @ e_i]$, and since this is nonstative,

$$(\exists e: [e \text{ recent-subep } e_i] [[\text{John wake-up}] * e]).$$

Since further $[[\text{John wake-up}] ** e_1]$ entails $[[\text{John awake}] * (\text{fin } e_1)]$ for all e_1 (via suitable meaning postulates), this suggests

$$[[\text{John awake}] * (\text{fin } e_i)],$$

where as in (13), e_i lasts until η , the (present) reference episode. In other words, John is still awake. Note that this is *only* an implication, since we can perfectly well say, without contradiction, “John has woken up and fallen asleep again.”

The second variant of perfect is simpler, amounting to a “relative past”:

$$\text{Perf}(2): [(\text{perf}_2 \Phi)_{\mathbf{T}} ** \eta] \\ \leftrightarrow [(\exists e_{\mathbf{T}}: [[e_{\mathbf{T}} \text{ before } \eta] \wedge [\text{Last}_{\mathbf{T}} \text{ orients } e_{\mathbf{T}}] \\ [\Phi_{\mathbf{O},\mathbf{T}} ** e_{\mathbf{T}}]]) ** (F \eta)]$$

Note that in the *Perf*(1)-rule, episode $e_{\mathbf{T}}$ was taken to last *until* the reference episode, and to be only *indirectly* characterized by Φ (via “@”); here $e_{\mathbf{T}}$ is *before* the reference episode, and is directly characterized by Φ (after de-indexicalization). The two main consequences are, first, that there is no longer an implication that a stative episode persists to the reference time; and second, the episode stored in the tense tree is now the actual Φ -event, not an episode which by inference *contains* a Φ -event. The differences between *perf*₁ and *perf*₂ help to account for the following contrasts:

- (15) a. *John has left yesterday.
 b. John had left the day before.
 c. John will have left the day before.
 (16) a. John has woken up and *has immediately dressed.
 b. John had woken up and had immediately dressed.
 c. John will have woken up and will immediately have dressed.

On our account, only *perf*₁ is available in the a-sentences; so in (15) a, we do not get “John’s leaving”

stored in the tense tree, but only an episode *containing* John’s leaving and lasting till now – but that cannot possibly be contained in *yesterday*. Similarly in (16) a, John’s waking up is not directly stored in the tense tree, and so is not directly available in the interpretation of *immediately*. (We think that there may be a recovery mechanism in the interpretation of adverbials, which looks for suitable points of orientation in the *history list* (where we assume that *all* indefinites and definites are recorded, whether explicitly mentioned or inferred) after failing to find them in the tense tree.) In the b and c sentences in (15) and (16), no difficulties are encountered because of the availability of the *perf*₂ reading.

Perhaps the most important point about our treatment of the perfect, alluded to earlier, is that the reference episode is introduced simply by the normal effect of operators (usually *past*, *pres* or *futr*) “exterior” to it. Briefly reconsidering (9) b, for example, we see that the (wide-scope) past will generate an episode in the past relative to the time of speech, and oriented by the “entering” episode in (9) a. It is this past episode which becomes the reference episode for the perfect. Now the key to the seemingly anaphoric character of this reference episode is this: if $[e_1 \text{ orients } e_2]$, and e_2 is *stative*, then there is a strong suggestion that e_2 is either concurrent with e_1 (namely, when e_1 is stative as well) or contiguous with its end point (when e_1 is nonstative). This is apparent, for instance, in the following variant of (9):

- (17) a. John entered the room.
 b. His paintings were gone.
 c. In their place were Schwarzenegger posters.

In b, the episode of the paintings being gone covers (at least) a short time-span immediately following John’s entering. In c, the episode of the posters being in place is concurrent with the paintings being gone, since both episodes are stative. Now, recognizing that perfect reference episodes are stative (because the state of being *after* some given type of event can persist indefinitely, and holds of the subintervals of any intervals of which it holds), the same analysis places the past episode in (9) b (serving as reference episode for the perfect) right after John’s entering, contiguous with it. This is tantamount to making the end of John’s entering the “reference time” for the perfect. So we get the desired “anaphoric” effect, *without* treating past as anaphoric, and without singling out the past in past perfect for special treatment.

Next we mention the de-indexicalization rules for *That*- and *that*-nominalizations:

$$\text{That}(1): (\text{That } \Phi)_{\mathbf{T}} = (\text{That } \Phi_{\rightarrow, \mathbf{T}}), \\ \text{tree transformation:} \\ (\text{That } \Phi) \cdot \mathbf{T} = \leftarrow (\Phi \cdot (\mapsto \mathbf{T}))$$

$$\text{That}(2): (\text{that } \Phi)_{\mathbf{T}} = (\text{That } \Phi_{\leftarrow, \mathbf{T}}), \\ \text{tree transformation:} \\ (\text{that } \Phi) \cdot \mathbf{T} = \leftarrow (\Phi \cdot (\leftarrow \mathbf{T}))$$

The rules cause focus shifts to the root of an *embedded* tree, indicated by the \mapsto and \leftarrow operations. The

first of these always adds a *new* embedding link whose destination is a *new* root node, whereas the second only does so if no embedding link exists at the current focus; otherwise, it causes *re-traversal* of the *last* embedding link added at the current focus. The leftward arrows in the two rules indicate focus restoration much like the upward arrows in the preceding rules. Intuitively, the distinction between a sentence nominalization derived explicitly from the text and one introduced through an implicit performative is motivated by the following sort of contrast:

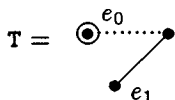
- (18) a. John knows that Mary got home before midnight.
 b. He also knows that she watched a movie.
 (19) a. Mary got home before midnight.
 b. She watched a movie.

In (18) the sentences about Mary in **a** and **b** are objects of attitudes, and it is much less clear than in (19) whether they refer to *successive* episodes. Now, the LFs for (19) **a**, **b** will contain performative predicates and *that*-nominalization operators; since the rule for *that* causes re-traversal of embedding links, the **a** and **b** episodes will end up at the same node. But in (18) only the “knowing” episodes map to the same node, while the episodes involving Mary map to different embedded nodes. Thus it will be harder to establish a connection between them (i.e., tense structure alone will not establish a connection though inference based on other information still might).

Before proceeding to further rules we illustrate some of the ones so far with a “trace” of an LF de-indexicalization, namely, that of sentence (21), uttered right after (20):

- (20) Mary looked pale.
 (21) John realized that Mary was tired.

The tree structure after processing of (20) will be



where e_0 corresponds to the added performative – i.e., the speaker’s utterance of (20), and e_1 corresponds to Mary’s looking pale at a point in the past. The logical form of (21) is initially (without performative)

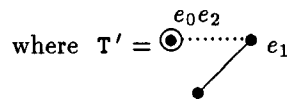
[John <past realize> (That [Mary <past tired>])].

and after scoping and addition of the implicit performative,

(pres [Speaker assert (that (past [John realize (That (past [Mary tired]))]))]).

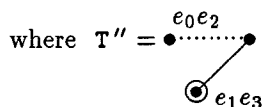
By the **Pres**-rule, the first step of the de-indexicalization relative to T yields

$(\exists e_2: [[e_2 \text{ at-about } Now_2] \wedge [e_0 \text{ orients } e_2]]$
 [[Speaker assert (that (past [John realize (That (past [Mary tired]))]))]] $_{T'} ** e_2)$

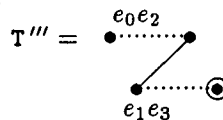


Next, a simple rule we have not mentioned moves the T' inward to the *that*-clause. The **That(2)**-rule causes retraversal of the embedding link, and the **Past**-rule then gives the following altered form of the outer *that*-clause:

(That $(\exists e_3: [[e_3 \text{ before } e_2] \wedge [e_1 \text{ orients } e_3]]$
 [[John realize (That
 (past [Mary tired]))]] $_{T''} ** e_3)$)



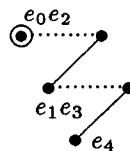
Again T'' is moved inward to the embedded *That*-clause, and the **That(1)**-rule generates a new embedding link and tree root at the focal node. It then remains to process the innermost tensed clause (past [Mary tired]) $_{T'''}$ with



One more application of the **Past**-rule, but keeping in mind that we are now at a past-dominated node, converts the tensed clause to

$(\exists e_4: [e_4 \text{ at-or-before } e_3] [[\text{Mary tired}] ** e_4])$

assuming the orienting predication is omitted when there is no orienting episode. The final tree structure will be



Since e_4 is stative (given its characterization [Mary tired]), the inference from the context-charged predication is that e_4 is concurrent with e_3 (i.e., John’s realization), in the absence of contrary information. Also the earlier context-charged relation $[e_1 \text{ orients } e_3]$ will lead to the inference that John’s realization was *during* Mary’s looking pale, in view of the fact that e_1 is stative and e_3 nonstative. (As well, a causal relation can be tentatively inferred.) The results of de-indexicalization thus seem to be in complete accord with intuition.

Besides the above rules, our current theory includes rules for conjunction, negation, quantification, and some adverbials and nominalizations. However, space limitations prevent inclusion of any details. Untensed conjunctions, negation and quantification introduce episodes for the clauses they operate on, explaining phenomenon like illustrated in (3) and (4). Adverbs and PP-adverbials of temporal location, duration, and manner are treated in a uniform way that de-indexicalizes them into predications about episodes. However, we make no attempt as yet to

interpret anaphoric NPs in those PPs (such as the NP in *before the war*). Also, relative clauses and clausal adverbials remain largely beyond the scope of our present work.

V. Conclusions and Further Work

We have described a new, principled approach to tense and aspect interpretation within a compositional framework for language understanding. The central concept is that of a tense tree structure as part of a more general context structure. This provides a straightforward and easily visualized way of converting originally indexical LFs to representations of the meaning of an utterance which are context-independent at least as far as the episodic relations implicit in the tense-aspect structure are concerned. This includes the most common “orienting” relations between episodes, and some of the more subtle relations conveyed by perfective aspect.

Our Common Lisp implementation of the de-indexicalization process allows rules to be straightforwardly represented and easily edited; example sentences of the type in this paper (*modulo* simplification of some of the adverbials) run in roughly a tenth of a second on a Sun SPARCstation 1.

Future work will focus on extension of the rules so as to cover more types of adverbials (especially clausal ones), tenseless VPs and clauses (some of which we already handle, such as action nominalizations like “To have loved and lost is not unusual”), and relative clause. A preliminary look at some of these phenomena, especially clausal adverbials and relative clauses, suggests that an operation on *pairs* of trees may be required, something one might call a “two-point overlay” $T_1:T_2$, where the root node of T_1 and another node of T_1 are aligned with two nodes of T_2 , creating orienting relationships between certain episodes stored at the aligned nodes. In any case, the potential of our approach has certainly not been exhausted.

Acknowledgements

We are grateful for some interesting suggestions from Bob Wilensky concerning the semantics of perfect, which caused us to modify our rules. The research was supported by NSERC Operating Grant A8818 (LKS), an Izaak W. Killam Memorial Scholarship (CHH), the Boeing Co. under Purchase Contract W-288104, and ONR/DARPA research contract no. N00014-82-K-0193.

References

- [1] Allen, J. *Natural Language Understanding*. Benjamin/Cummings Publ. Co., Reading, Mass., 1987.
- [2] Barwise, J. *The Situation in Logic*. CSLI: Stanford, CA, 1989.
- [3] Davidson, D. “The logical form of action sentences.” In D. Davidson and G. Harman, eds., *The Logic of Grammar*, 235–245, Dickenson, Encino, CA, 1975.
- [4] Dowty, D. “Tense, time adverbs and compositional semantic theory.” *Ling. and Phil.*, 5:23–55, 1982.
- [5] Harman, G. “Logical form.” In D. Davidson and G. Harman, eds., *The Logic of Grammar*, 289–307, Dickenson, Encino, CA, 1975.
- [6] Hinrichs, E. W.. “Tense, quantifiers, and contexts.” *Comput. Ling.*, 14(2):3–14, 1988.
- [7] Hobbs, J. R., “Coherence and coreference”, *Cog. Sci.*, 3(1):67–82, 1979.
- [8] Leech, G. *Meaning and the English Verb* (2nd edition), Longman, London, 1987.
- [9] McCawley, J.D., “Notes on the English present perfect”, *Australian J. of Ling.*, 1:81–90, 1981.
- [10] Moens, M. and Steedman, M. “Temporal ontology and temporal reference.” *Comput. Ling.*, 14(2):15–28, 1988.
- [11] Passonneau, R.J., “A computational model of the semantics of tense and aspect.” *Comput. Ling.* 14(2):44–60, 1988.
- [12] Partee, B., “Some structural analogies between tenses and pronouns in English.” *J. of Phil.*, 70: 601–609, 1973.
- [13] Reichenbach, H. *Elements of Symbolic Logic*. Macmillan, New York, NY, 1947.
- [14] Richards, B., “Tenses, temporal quantifiers and semantic innocence”, in E. LePore (ed.), *New Directions in Semantics*, 337–384, Academic Press, New York, NY, 1987.
- [15] Schubert, L. K. and Hwang, C. H. “An episodic knowledge representation for narrative texts.” In *1st Inter. Conf. on Principles of Knowledge Representation and Reasoning (KR89)*, 444–458, Toronto, Canada, May 15–18, 1989.
- [16] Schubert, L. K. and Hwang, C. H. “An episodic knowledge representation for narrative texts.” TR 345, U. of Rochester, Rochester, NY, May 1990.
- [17] Schubert, L. K. and Pelletier, F. J. “From English to logic: context free computation of ‘conventional’ logical translations.” *American J. of Comp. Ling.*, 8:26–44, 1982. Also in *Readings in Natural Language Processing*, B. Grosz, K. Jones and B. Webber, eds., 293–311, Morgan Kaufman, Los Altos, CA, 1986.
- [18] Schubert, L. K. and Pelletier, F. J. “Generically speaking, or, using discourse representation theory to interpret generics.” In G. Chierchia, B. Partee, and R. Turner, editors, *Property Theory, Type Theory, and Semantics, V.2: Semantic Issues*, 193–268. Kluwer Academic Publ., Boston, MA, 1989.
- [19] Webber, B. L. “Tense as discourse anaphor.” *Comput. Ling.*, 14(2):61–73, 1988.