TEMPORAL CENTERING

Megumi Kameyama
SRI International AI Center
333 Ravenswood Ave.,
Menlo Park, CA 94025
megumi@ai.sri.com

Rebecca Passonneau
Dept. of Computer Science
Columbia University
New York, NY 10027
becky@cs.columbia.edu

Massimo Poesio
Dept. of Computer Science
University of Rochester
Rochester, NY 14627-0226
poesio@cs.rochester.edu

Abstract

We present a semantic and pragmatic account of the anaphoric properties of past and perfect that improves on previous work by integrating discourse structure, aspectual type, surface structure and commonsense knowledge. A novel aspect of our account is that we distinguish between two kinds of temporal intervals in the interpretation of temporal operators — discourse reference intervals and event intervals. This distinction makes it possible to develop an analogy between centering and temporal centering, which operates on discourse reference intervals. Our *temporal property-sharing principle* is a defeasible inference rule on the logical form. Along with lexical and causal reasoning, it plays a role in incrementally resolving underspecified aspects of the event structure representation of an utterance against the current context.

The Problem

The past tense has been compared with anaphoric definite pronouns ([20] [22]) and definite noun phrases ([27]). The supporting observation is that in two consecutive past tense descriptions of events, as in (1) below (from [18]), the second sentence refers to a time (t') whose identity depends on the time (t) of the event described in the first sentence.

(1)a. The Lone Ranger got on his horse. (t)
    b. He rode off into the sunset. (t')

Tense interpretation also involves commonsense inferences in that the specific relation between the two event times may vary. In (1), the relation inferred to hold is temporal progression (t $\prec$ t'), but other ordering relations are also possible (see [6] [27]). Any temporal relations are in fact possible for two consecutively described events in discourse.

A number of factors affect the interpretation as to whether successive past tenses are anaphorically related, and if they are, what the relative order of the associated events is. The determinant factors have been argued to be *discourse structure* ([27] [14]), *aspectual type* ([6] [12] [17]), *surface structure* ([7] [14]), and *commonsense knowledge* ([19] [27] [13]). However, no account has adequately addressed all four factors.

The problem in tense interpretation that we address is illustrated with Example (2) (from [27]).

(2)a. John went over (t1) to Mary's house.
    b. On the way, he had (t2) stopped (t3) by the flower shop for some roses. (t3 $\prec$ t2 (=t1))
    c. Unfortunately, they failed (t4) to cheer her up. (t3 $\prec$ t1 $\prec$ t4)
    c'. He picked out (t4') 5 red ones, 3 white ones and 1 pale pink. (t3 $\prec$ t4' $\prec$ t1)

(2c) and (2c') are alternative third sentences. Although both are in the simple past, and both evoke events of the same aspectual type (transition event [23]), they are interpreted differently. We refer to the contextually established time that a past tense is resolved against as the "discourse reference time." A discourse reference time (t1) is introduced in (2a) with the event of John going to Mary's house at t1.[1] The past perfect in (2b) introduces two times: John's stopping at the flower shop (t3) *precedes* the time t2 (t3 $\prec$ t2), and t2 is typically inferred to be equal to the time of going over to Mary's house (t1); hence t3 $\prec$ t1. In (2c), the time of failing to cheer Mary (t4) is inferred to occur just *after* t1, whereas in the parallel version (2c'), the time of picking out roses (t4') is inferred to occur *during* or just *after* t3, thus *before* t1. Thus, as noted by Webber [27], a past perfect evokes two temporal referents (e.g., t2 and t3) and either of them can provide a discourse reference time for a subsequent past tense.

---

[1]Later we will see that although the first tense generates a discourse reference time for the subsequent one, its temporal referent (here t1) is not precisely equivalent to it.

Another aspect of the problem in tense interpretation is illustrated with so-called *extended flashbacks* such as Example (3), similar to the one discussed in [14].

> (3) John and Mary *went* to buy a lawnmower. Somebody *had stolen* theirs the day before. They *had seen* the thief go away. John *had run* after him to no avail. All the lawnmowers *were* too expensive. They *decided* they couldn't afford a new one.

There are two narrative threads in (3), one describing John and Mary's visit to a store, the other John's chase of the thief. These threads correspond to the repeated use of the past and of the past perfect, respectively. The return from the past perfect to the past in also coincides with the return to the earlier thread. Note that (i) not only the simple past but also the past perfect can achieve an effect of continuity and that (ii) more than one discourse reference time needs to be maintained in order to account for the return to an earlier one.

The general problem in tense interpretation that we address is how to keep track of all the potential discourse reference times in a context, and how to select the right one for a given anaphoric past tense.

We argue that the choice of the discourse antecedent of a temporal operator is subject to *centering effects*. We assume that each temporal operator in a sentence introduces a discourse reference time into the discourse context. We claim that this set of times constitutes a list of potential discourse reference times for the next sentence, which we'll later refer to as the temporal forward-looking center (TCf), and that the position of a temporal operator in the logical form of the sentence affects the choice of the antecedent through structural parallelism (as a case of the *property-sharing* effect in centering [16]).

We formalize the effect of surface structure on the choice of temporal antecedent by means of defeasible axioms. These axioms must be *less specific* than axioms encoding causal reasoning. We argue that the choice of discourse reference time is an instance of a general principle in defeasible reasoning, namely, the *Penguin Principle* [19] that chooses the *most specific* axiom applicable. We support our claims with data from the Brown corpus.

In the next section, we review the three existing proposals most related to ours — Webber [27], Lascarides and Oberlander [19], and Hwang and Schubert [14]. The next two sections present a review of centering followed by a discussion of the analogous effects of temporal centering. To account for temporal centering effects, we then present our tense rules and our account of how commonsense reasoning interacts with the default preferences generated by the surface structure.

Related Work

Webber [27] argues that temporal entities are subject to focusing processes analogous to those constraining the interpretation of definite noun phrases. She explicitly rejects, however, a more direct analogy to Sidner's [26] potential local foci, and assumes only one temporal referent in the *temporal focus* (TF).

Lascarides and Oberlander [19] present defeasible reasoning rules for narrative understanding that partly address the reasoning and control needs of Webber's model. For example, they argue that in the case of conflicting inferences regarding the temporal order of two times [e.g., whether precedes, follows, or overlaps], the *most specific* interpretation should be preferred. However, they do not address cases such as (2), where there are multiple possible discourse reference times and the choice of one interpretation over the other needs to be accounted for.

Hwang and Schubert [14] intend to account for the role played by structural factors in the choice of anaphoric relations among distinct past tenses. They propose the contextual structures called *tense trees* built as a narrative gets interpreted. Tense trees reflect the structural dependencies among the tense and aspect operators in the interpretation of the sentences. The events evoked by a sentence are "appended" to the tree nodes whose positions reflect the structural positions of the corresponding temporal operators in the sentence.

Webber [27] and Hwang and Schubert [14] differ markedly regarding the relative weight they assign to commonsense inferences versus surface structure, but both have problems with examples like (2). Neither can readily account for the interpretation assigned to (2c'). Webber assumes that the TF of (2b) is initially t1, but shifts to t3, thus initiating an embedded segment, *after* (2c') has been processed sufficiently to recognize that TF=t3 yields a *more sensible* interpretation than TF=t2(=t1). She does not discuss how t3 comes to be considered as a potential TF in this case, much less how to reject t2(=t1). Like Webber, Hwang and Schubert assume that (2c') requires recognition of an embedded segment, but admittedly they have not resolved the conflict between the need for discourse segment recognition, a task for commonsense reasoning, and the clear separation of tense resolution from commonsense reasoning that they advocate. They also fail to distinguish betwen the hierarchical structures of sen-

tences and discourses in the tense tree. Tense tree thus overcommits in places where the rule should belong to defeasible preferences.

Our approach is to formulate the structural preferences for tense resolution as defeasible rules analogous to centering preferences for pronoun resolution. These structural preferences, which are overlooked in Webber's or Lascarides and Oberlander's accounts, interact with the commonsense inferences used in tense interpretation.

Centering

Centering [9] is a refinement of Sidner's [26] local focusing model. It consists of a set of principles and rules for dynamically updating the local *attentional state* [11] in discourse processing.

We assume the following general picture of discourse processing. A discourse consists of a sequence of utterances $utt_1, ..., utt_n$. The sentence grammar translates the content of each utterance utt_i into a (set of) *surface* logical form(s) containing unresolved anaphoric expressions and operators. We call it here a "surface" formula ϕ_i. This logical form is similar, in spirit, to Hwang and Schubert's [14] *indexical formula* and Alshawi's [2] *quasi logical form*, whose main motivations are to represent that part of the sentence meaning independent from the particular discourse context. This "baseline" meaning representation acts as a clean interface to the pragmatic processing needed to resolve context-dependent expressions. Utterance interpretation takes place in a context, and outputs an updated context. Part of this dynamic context is the attentional state that represents the *currently salient entities* partially ordered by relative salience.

We say that each formula ϕ_i defines a *transition relation* between the input attentional state AS_{i-1} and the output attentional state AS_i. An attentional state AS_i contains the *focus of attention* Foc_i whose most salient subpart is the *center of attention*. After resolving anaphoric expressions in ϕ_i, the center of attention contains a partially ordered set of *forward-looking centers* Cf_i comprising the entities realized in ϕ_i. A member of Cf_i might (but need not) be the *backward-looking center* Cb_i, the currently most salient entity.

Centering has mainly been used to constrain how discourse anaphoric pronouns are processed; e.g., the centering rule [9] predicts that Cb_i will be realized with a pronoun if $Cb_i=Cb_{i-1}$.[2] Also, when $Cb_i=Cb_{i-1}$ and both are realized by definite pronouns, it is predicted that both will be real-

a.	John went to the store.	
	Cf_1=[John',store1]	Cb_1=NULL
b.	He saw Bill.	
	Cf_2=[John',Bill']	Cb_2=John'
	Cb-establishment	
c.	He walked towards him.	
	Cf_3=[John',Bill']	Cb_3=John'
	Cb-retention	
c'.	He appeared pale to him.	
	$Cf_{3'}$=[Bill', John']	$Cb_{3'}$=Bill'
	Cb-establishment	

Figure 1: Illustration of Centering

ized in a common grammatical role, with subject preferred over non-subject [16] [24]. A number of transition relation types have been distinguished in centering. In this paper, we will use the following four types: *Cb-retention, Cb-establishment, Cb-resumption*, and *NULL-transition*.[3]

In Cb-retention, the same entity is retained as the *Cb*: $Cb_{i-1} = Cb_i \neq NULL$. In Cb-establishment, another member of Cf becomes the Cb: $Cb_{i-1} \neq Cb_i$, $Cb_i \in Cf_{i-1}$ (Cb_{i-1} may be NULL). These two are the most relevant to the present paper. In Cb-resumption, an old *Cb* not in Cf but in the current *Foc* is resumed. In NULL-transition, the output state has no Cb (Cb_i=NULL). Centering posits a *default preference* for retention over establishment. We provisionally assume that establishment is preferred over resumption or NULL-transition.

We illustrate centering with Fig. 1, where c and c' are alternative continuations of b. After a., Cf_1 contains two entities, John' and store1. In b., John' is referred to with a subject pronoun, and is established as Cb_2. In c., because John' is the current Cb, and because retention is preferred over establishment, centering predicts that a subject pronoun will refer to John' rather than to Bill'. The default is overridden in c' and instead, the subject pronoun is inferred to refer to Bill' because it is likely that the perceiver in the first perceptual state, see', remains the perceiver in the subsequent perceptual state, appear'.

Figure 2: Illustration of Temporal Centering

Centering effects in tense interpretation

Our proposal is motivated by the parallel between example (2) above and the centering example in Fig. 1. As illustrated in Figure 2, we would like to say that utterances (2a) and (2b) share the same discourse reference time r1 that links t2 to t1. The shared discourse reference time r1 is thereby established as a backward-looking temporal center (TCb). (2c) retains the TCb, linking t4 to t2(=t1), whereas (2c') establishes a new TCb, linking t4' to t3.

In order to establish a direct analogy between centering and temporal centering, however, we need to first clarify the nature of the temporal entities in the attentional state. Note that if (2c) retains the "same" TCb, this TCb cannot be identified with either t2 of (2b) or t4 of (2c), since t2 and t4 are distinct (t1=t2, t1 $\prec$ t4). The TCb remains fixed while the event times within it may move forward or backward. The TCb is then quite different from the *reference time* as used in proposals inspired by Reichenbach [25] such as Dowty's [6] and Hinrichs [12]).[4] Recall the extended flashback example (3) above. There are two simultaneous narrative threads, in the simple past and in the past perfect, and the former seems to remain somehow in the background while the latter is in the foreground. These examples seem to indicate that when processing a text a reader maintains a list of "potential foci" ordered by their relative salience, instead only one temporal focus, as in Webber's account.

Further evidence in favor of the analogy was obtained by examining a random sample of dis-

---

[4] A similar proposal is made by Kamp and Reyle [17], where they argue for the separation of the *temporal perspective point* (*TPpt*) that is relatively fixed and the *reference point* (*Rpt*) that moves with narrative progression.

course sequences from the Brown corpus [8], a heterogeneous corpus that should yield unbiased data. Each multi-sentence sequence contained one of two types of *trigger sentences* with multiple temporal operators — the past perfect matrix clause or the past matrix verb taking an embedded past or past perfect complement clause (e.g., *John remembered that Mary had asked for roses*). We observed that this trigger's output TCb was almost always the time associated with the superordinate operator. Following each trigger was a *target sentence* with a simple past matrix clause. This creates a sequence analogous to (2b-c) or (2b-c'). We analyzed each sequence to determine whether the discourse reference time for the past in the target sentence was introduced by the trigger's superordinate operator (as a case of TCb-retention) or by a subordinate operator (as a case of TCb-establishment). In the total of 80 examples, we found 52 retentions (65%) and 12 establishments 15%). Others included 9 cases of "neither" (11%) (some combination of resumption and NULL-transition) and 7 undecidable cases (9%). These data are consistent with our hypothesis that in a sentence with multiple temporal operators, the relative likelihood that a given temporal operator provides the discourse reference time for a subsequent past depends on its syntactic prominence. The question is, how do temporal centering and commonsense reasoning interact to yield the observed results?

Two levels of logical representation

Our explanation for the interaction between temporal centering and commonsense reasoning rests on assuming two distinct levels of representation in sentence interpretation. One is the *logical form* independent from the context and obtained by a direct translation of the surface syntactic structure of the utterance. The other is a fully resolved logical expression that results from incrementally resolving context-dependent expressions in the logical form. For simplicity, our discussion ignores anything but temporal operators.

Much as in Hwang and Schubert's proposal [14], the logical form contains unresolved temporal operators — e.g., tense operators, PRES(ent) and PAST, and aspectual operators, PERF(ect) and PROG(ressive). It also represents the structural position of the temporal operators in a sentence. The crucial difference is that we take each tense and aspect operator to also give rise to a *discourse reference interval* (see below) that is contextually resolved. Our logical forms for (2a) and (2b) are shown in (4).

(4)a. $(\text{PAST}_{r1}$ 'John goes over to Mary's house')

73

```
τ(PAST_rφ)=(∃ e (e ⊆ r) ∧
                (e starts_before SuperNow) ∧
                (LastIn(r) orients e)
             [e ⊨ τ(φ)])
τ(PERF_rφ)=(∃ e (e ⊆ r) ∧
                (e starts_before SuperRef) ∧
                (LastIn(r) orients e)
             [e ⊨ τ(φ)])
```

Figure 3: Mapping rules for PAST and PERF

b. (PAST_r2 (PERF_r3 'John stops by the florist for some roses'))

The temporal operators in our logical forms are translated into the language of what we call the *event structure representation (ESR)*. In essence, ESR represents the temporal and causal relations among the eventualities described in discourse. We will use Allen's [1] interval representation scheme for this representation level. Described eventualities correspond to *event intervals* and what we have been calling the "discourse reference times" correspond to *discourse reference intervals*. To represent relations among these temporal intervals, we use disjunctive subsets of Allen's thirteen relations. With Allen's representation, we can directly represent vague or ambiguous temporal relations expressed in natural language and incrementally resolve them to more specific relations using constraint propagation. Our discourse reference intervals coincide exactly with the computational role of Allen's *reference intervals*.

Tense mapping rules

Now we define the recursive mapping τ between the logical form and the event structure representation.[5] These tense mapping rules "unpack" the relations among relevant event intervals and discourse reference intervals encoded by temporal operators in the logical form, and generate the initial event structure representation. Although these rules look similar to Hwang and Schubert's tense deindexing rules, they play a very different role. Rather than performing sentence interpretation, as their deindexing rules do, our mapping rules specify the semantics of the logical form in terms of the event structure representation. The τ rules for PAST and PERF are shown in Fig. 3.

We assume that formulas denote sets of time intervals. For every clause associated with a formula OPφ, where OP is a temporal operator (PAST or PERF) and φ is a formula, there is an event inter-

val e corresponding to the specific time for which φ is inferred to hold. We express this relation as e ⊨ φ (e supports φ). A statement of this form evaluates to true if the denotation of e is an element of the denotation of φ.

Each event interval e is temporally included in a discourse reference interval r, semantically treated as an open variable — the value of r is constrained by contextual factors such as temporal centering and commonsense reasoning, as described in a later section.

SuperNow and SuperRef have values determined by the position of the subformula OPφ in the logical form. SuperNow evaluates to the *current utterance interval* — the time interval in which the current utterance takes place. We assume that there is an updating mechanism for SuperNow defined for each utterance initiation including the onset of a quoted speech. SuperRef evaluates to the event interval evoked by the temporal operator that immediately dominates the current one in the logical form.

Starts_before, a disjunction of meets, precedes and the inverse of finishes (m ≺ fi), is the possible relation between e and SuperNow for any PASTφ (or between e and SuperRef for any PERFφ), irrespective of the aspectual type of φ. Orients, borrowed from Hwang and Schubert, is the disjunctive set of all of Allen's thirteen relations. Both starts_before and orients may be further specialized after computing φ's aspectual type, which depends partly on the aspectual types of its constituents and partly on commonsense reasoning.[6] We can state certain default specifications of orients. For example, involving two transition event intervals, orients defaults to *precedes* (e.g., see (1) where t ≺ t'). Stative intervals by default overlap other non-stative or stative intervals (cf. [6] [12]).

LastIn(r) is a function defined over discourse reference intervals, and evaluates to the most recent *non-stative* event interval in the discourse reference interval r, where applicable. This event interval orients the new event interval e. LastIn(r) corresponds to the "reference time" that moves with narrative progression in approaches like Dowty [6] and Hinrichs [12].

How the mapping rules work

Table 1 shows the logical form and the event structure determined by the mapping rules for each sentence of (2) (repeated here).

(5)a. John went over (t1) to Mary's house.

---

[5] The use of τ was inspired by the τ translation in temporal logic [4].

[6] See [23] for a partial model of computing aspectual type; see [21] for a partial model of aspectual coercion, one type of commonsense reasoning.

74

	ϕ_i	$e \subseteq r$	Constraint
a	PAST α	t1 $\subseteq$ r1	t1 ($\prec$) u_a
b	PAST		
	(PERFβ)	t2 $\subseteq$ r2	t2 (m fi) u_b
	PERF β	t3 $\subseteq$ r3	t3 ($\prec$) t2
c	PAST γ	t4 $\subseteq$ r4	t4 ($\prec$) u_c
c'	PAST γ'	t4' $\subseteq$ r4'	t4' ($\prec$) $u_{c'}$

Table 1: Application of Rules to (2)

b. On the way, he had (t2) stopped (t3) by the flower shop for some roses.

c. Unfortunately, they failed (t4) to cheer her up.

c'. He picked out (t4') 5 red ones, 3 white ones and 1 pale pink.

The symbols α–γ' correspond to the tenseless propositions of sentences a–c', and u_a––$u_{c'}$ represent the respective utterance event intervals. We explain the interpretation of sentence (2b) ((5b) above) in detail. Applying the rule for PAST, we obtain t2 $\models$ τ(PERFβ), with discourse reference interval r2 such that t2 $\subseteq$ r2. SuperNow evaluates to u_b, yielding t2 starts_before u_b. PERFβ is necessarily stative, so by default its event interval is assumed to persist up through the present. Thus t2 starts_before u_b is specialized to t2 (m fi) u_b. Applying the PERF rule, we obtain t3 $\models$ β, with discourse reference interval r3 such that t3 $\subseteq$ r3. SuperRef evaluates to t2, the event interval of the dominating PAST operator, yielding t3 starts_before t2, which is then specialized to t3 $\prec$ t2 due to the aspectual types of PERFβ and β.

The interaction between structural factors and causal knowledge

We propose that defeasible rules producing hypotheses about the interpretation of an utterance operate on both the levels of representation used in tense interpretation. On the event structure representation level, we assume the existence of axioms formalizing lexical semantics and other aspects of commonsense knowledge (e.g., [13] [19]), and call them *causal axioms*. Simplified examples are given in (6) and (7) below. We use the default conditional > introduced by Asher and Morreau [3] and used by Lascarides and Oberlander [19]. (6) encodes the inference that people usually pick up flowers at the florist's. (7) say that people are usually at certain places as a result of their going there.

(6) PICK-UP(X,flowers,t) >
    (AT(X,florist,t') $\wedge$ t$\subseteq$t')

(7) AT(X,place,t) >
    (GO-TO(X,place,t') $\wedge$ t'$\prec$t)

The other, more novel, assumption we make is that there are defeasible rules that operate on the logical form, on the basis of structural information only. We formalize the effects of structural information in terms of what we call *logical form axioms*. Much as in the case of axioms formalizing commonsense knowledge, logical form axioms can be defeasible. Although the effects of structural factors on anaphora interpretation are widely acknowledged in linguistic theories (e.g., binding theory), the interest is more on indefeasible grammatical principles. Our logical form axioms encode grammatically derived defeasible preferences, the kind of structural effects that centering and temporal centering uncover.

We assume the following architecture for sentence interpretation. Once the logical form has been obtained, logical form reasoning takes place. There are two kinds of rules in logical form reasoning — monotonic (or indefeasible) and defeasible. The former rewrite the logical form, and the latter generate *logical form hypotheses*, that are additional logical form expressions. All of the hypotheses resulting from logical form reasoning, as well as the original interpretation, are then mapped into event structure representations, and causal axioms are used to complete the interpretation. All complete interpretations of a sentence are compared to find the best hypothesis.

We propose to encode temporal centering as defeasible logical form axioms that generate logical form hypotheses about the interpretations of unresolved discourse reference intervals. To be more precise, we claim that the following principle holds:

Temporal Property-sharing Principle (**TPP**)
Resolve the current discourse reference interval r_i against a discourse reference interval r_j in the attentional state. By default, r_i and r_j should share the same *property* reflecting the temporal position in the logical form.

This general preference can be encoded by means of logical form axioms of the type shown in (8), whose intended interpretation is as follows — Let utt_{n-1} and utt_n be two consecutive utterances in discourse, lf(utt) be a function from utterances to their logical forms, (PAST$_{r1}$ Φ) be a subformula of lf(utt_{n-1}), (PAST$_{r2}$ Ψ) be a subformula of lf(utt_n), and pos(lf1,lf2) be a function from pairs of a formula lf1 and a subformula lf2 into the "temporal position" of lf2 in lf1. The *temporal position* of a subformula in a formula is the sequence of *higher* temporal operators of the subformula. For example, in the for-

mula (PAST (PERF ϕ)), the temporal position of (PAST (PERF ϕ)) is the empty sequence $\langle\rangle$, that of (PERF ϕ) is the sequence $\langle$PAST$\rangle$, and that of ϕ is the sequence $\langle$PAST,PERF$\rangle$. In a case where pos(u$_{n-1}$,(PAST$_{r1}$ Φ) = pos(u$_n$, (PAST$_{r2}$ Ψ)), the TPP generates the hypothesis that the two discourse reference intervals r1 and r2 are the same.

$$
(8) \quad \frac{
\begin{array}{l}
\texttt{consecutive(utt}_{n-1}\texttt{,utt}_n\texttt{)},\\
\texttt{pos(lf(utt}_{n-1}\texttt{),(PAST}_{r1}\ \Phi\texttt{))} =\\
\texttt{pos(lf(utt}_n\texttt{), (PAST}_{r2}\ \Psi\texttt{))}
\end{array}
}{
\hookrightarrow \\[2pt]
\texttt{utt}_n\texttt{:} \quad \texttt{(PAST}_{r2=r1}\ \Psi\texttt{)}
}
$$

A similar rule generates hypotheses about the identity of two discourse reference intervals both introduced by PERF operators in the same temporal positions.

Lascarides and Oberlander [19] propose the *Penguin Principle* as a way of choosing among conflicts generated by the application of defeasible rules. The principle says that whenever a conflict between rules arises, the hypothesis obtained by applying the *most specific* rule should be chosen. The logical form axioms encode general knowledge of linguistic structures independent from particular events or situations being described. Thus the information they encode is necessarily *less specific* than rules encoding lexical inferences and causal relations. Hence the following consequences on the TPP:

> The TPP applies in the absence of information about causal relation between the events. Any hypothesis that results from causal reasoning is preferred over the hypothesis suggested by the TPP.

This general pattern of interaction between causal and logical form axioms predicts the following Temporal Centering Principle:

Temporal Centering Principle (TCP) When ϕ_i evokes a set of unresolved discourse reference intervals $r_1,...,r_n$, each r_k is resolved against the current atttentional state as follows — (i) *IF* causal axioms generate a hypothesis, take it, (ii) *ELSE IF* the TPP generates a hypothesis, take it, (iii) *ELSE* $r_1,...,r_n$ are new to the discourse. After resolution, $r_1,...,r_n$ are in the TCf_i in the output context of ϕ_i.

Case (i) of the TCP allows any of the four transition types, retention, establishment, resumption, or NULL. A NULL-transition only result only from negative conclusions obtained from causal axioms. Under the TPP in the form of rule (8) that applies only on consecutive utterances, case (ii) would result in either TCb-retention or TCb-establishment, but not TCb-resumption.

How does this accord with our corpus data discussed earlier? In the most frequent (65%) TCb-retention, the discourse reference interval evoked by the matrix past tense is identified with another evoked by the matrix past tense in the previous utterance. In the next frequent (15%) TCb-establishment, it is identified with the one evoked by the subordinate temporal operator. This sharp difference in frequency would be explained by the fact that the TPP supports the former but not the latter.

Webber's example, revisited.

We now go back to Webber's example (2) and explain how the interactions of defeasible rules result in the temporal centering transitions shown in Fig. 2. The input consists of the logical forms from Table 1.

(2a) For the initial sentence of discourse, only the NULL-transition can apply. r1 is not identified, and remains as an open variable. AS_1 contains TCf_1=[r1].

(2b) The TPP can apply for PAST; resulting in TCb-establishment, with r1=r2. As a result, TCf_2=[r1,r3], TCb_2=r1, and the value of LastIn(r2)=LastIn(r1) is t1. This yields t1 orients t2, where t1 is the transition event interval for (α) and t2, for a state (PERF β). When a non-stative event interval orients a stative event interval, the former is *partly covered by* (overlaps, starts, finishes, equals, during, or their inverses) the latter by default.

(2c) The TPP, in the form of rule (8), generates the hypothesis expressed by the following logical form: (PAST$_{r4=r1}$ 'The roses fail to cheer her'). The causal reasoning generates the same conclusion, namely, *in order to give somebody flowers, one has to be at somebody's place*, so we conclude that the 'failing-to-cheer event' continues the "story" of John's going over to Mary's house. Either way, we have a TCb-retention, yielding TCf_3 =[r1(=r4)] and TCb_3=r1. The value of LastIn(r4) is now t1, yielding t1 orients t4. Since t1 and t4 are event intervals for transition events (α and γ), orients defaults to t1 ($\prec$) t4.

(2c') There is a conflict. While the TPP would suggest a continuation of the story of John's visit at Mary's house, the causal rules such as (6) suggest that the event of picking up flowers takes place at the florist's, and the rules such as (7) suggest that being at the florist's results from the action of "stopping by" the florist's. The Penguin Principle now applies, resulting in TCb-establishment, and t3 ($\prec$) t4.

Conclusions

We have presented a semantic and pragmatic account of the past tense that uniformly accounts for its discourse anaphoric properties. We distinguish between two kinds of intervals, discourse reference intervals and event intervals, which allows us to provide a direct analogy between centering and temporal centering.

We introduce the notion of logical form reasoning, which obviates the need for tense trees as the source for structural reasoning, and, at the same time, enables us to account for the interaction between the structural preferences and causal reasoning. By including in both logical form expressions and event structure representation an explicit indication of the contextual parameters to be resolved, rules can be formulated that exploit the kind of information available at each level. We can therefore factor out different sources of defeasible preferences in tense interpretation. Rules operating on the logical form formalize the effects of structural factors, and rules operating on the event structure formalize the effects of commonsense knowledge.

The ordering preferences in our model are consistent with data from a large, heterogeneous corpus. In the future work, we need to develop an evaluation method for our model, to specify more fully the knowledge representation and reasoning needs, and finally to clarify consequences that our model might have on a more unified model of attentional state to simultaneously account for nominal and temporal anaphora.

Acknowledgement

References

[1] James F. Allen. Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26:832–843, 1983.

[2] Alshawi, Hiyan. 1990. Resolving quasi logical forms. *Computational Linguistics.* 16(3), 133-144.

[3] Nicholas Asher and Michael Morreau. 1991. Commonsense entailment: a modal theory of nonmonotonic reasoning . In *Proceedings of the IJCAI 1991.*

[4] Johan van Benthem. forthcoming. Temporal logic. In D. Gabbay, Chr. Hogger and J. Robinson, eds. *Handbook of Logic in Artificial Intelligence and Logic Programming, Volume III.* (second revised version) Oxford University Press.

[5] Brennan, S., L. Friedman, and C. Pollard. 1987. A centering approach to pronouns. In *Proceedings of the 25th ACL Meeting,* 155-162.

[6] David Dowty. The effects of aspectual class on the temporal structure of discourse: Semantics or pragmatics. *Linguistics and Philosophy,* 9:37–61, 1986.

[7] Mürvet Enç. Anchoring conditions for tense. *Linguistic Inquiry,* 18.4:633–657, 1987.

[8] W. Francis and H. Kucera. *Frequency Analysis of English Usage: Lexicon and Grammar,* Houghton Mifflin, Boston, MA, 1982.

[9] Barbara J. Grosz, Aravind K. Joshi, and Scott Weinstein. Providing a unified account of definite noun phrases in discourse. In *Proceedings of the 21st ACL,* pages 44–50, 1983.

[10] Grosz, B., A. Joshi, and S. Weinstein. 1983. Towards a computational theory of discourse interpretation. Unpublished manuscript.

[11] Barbara J. Grosz and Candace L. Sidner. Attention, intentions and the structure of discourse. *Computational Linguistics,* 12:175–204, 1986.

[12] Erhard Hinrichs. Temporal anaphora in discourses of English. *Linguistics and Philosophy,* 9:63–82, 1986.

[13] Jerry R. Hobbs, Mark Stickel, Paul Martin, and Douglas Edwards. Interpretation as abduction. In *Proceedings of the 26th ACL,* pages 95–103, 1988.

[14] Chung Hee Hwang and Lehnart K. Schubert. Tense trees as the 'fine structure' of discourse. *Proceedings of the 30th ACL,* pages 232-240, 1992.

[15] Megumi Kameyama. *Zero Anaphora: The Case of Japanese.* PhD thesis, Stanford University, 1985.

[16] Megumi Kameyama. A property-sharing constraint in centering. In *Proceedings of the 24th Annual Meeting of the ACL,* pages 200–206, New York, 1986.

[17] Hans Kamp and Uwe Reyle. *From Discourse to Logic, Vol. I.* Kluwer.

[18] Lauri Karttunen. Discourse referents. In J. McCawley, editor, *Syntax and Semantics, Vol. 7: Notes from the Linguistic Underground.* Academic Press, New York, 1976.

[19] Alex Lascarides and Jon Oberlander. Temporal coherence and defeasible knowledge. *Theoretical Linguistics,* 1992. To appear.

[20] James D. McCawley. Tense and time reference in English. In Charles J. Fillmore and D. Terence Langendoen, editors, *Studies in Linguistic Semantics.* Holt, Rinehart and Winston, Inc., New York, 1971, pages 97-114.

[21] Mark Moens and Mark J. Steedman. Temporal ontology and temporal reference. *Computational Linguistics,* 14:15–28, 1988.

[22] Barbara Hall Partee. Some structural analogies between tenses and pronouns in English. *Journal of Philosophy,* 70:601-609, 1973.

[23] Rebecca J. Passonneau. A computational model of the semantics of tense and aspect. *Computational Linguistics,* 14:44–60, 1988.

[24] Rebecca J. Passonneau. Getting and keeping the center of attention. In R. Weischedel and M. Bates, editors, *Challenges in Natural Language Processing.* Cambridge University Press, To appear. Also available as Tech. Report CUCS-060-90, Dept. of Computer Science, Columbia University.

[25] Hans Reichenbach. 1947. *Elements of Symbolic Logic.* Macmillan Company. Reprinted, 1980, Dover Publications, New York.

[26] Candace L. Sidner. Focusing in the comprehension of definite anaphora. In M. Brady and R. C. Berwick, editors, *Computational Models of Discourse,* pages 267–330. The MIT Press, Cambridge, Massachusetts, 1983.

[27] Bonnie Lynn Webber. Tense as discourse anaphor. *Computational Linguistics,* pages 113–122, 1988.