MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A.I. Memo No. 1122 September, 1989

CAUSAL/TEMPORAL CONNECTIVES: SYNTAX AND LEXICON

Michael Brent

Abstract: This report elucidates the linguistic representation of temporal relations among events. It does so by examining sentences that contain two clauses connected words like *once, by the time, when,* and *before.* Specifically, the data involve the effect of the tenses of the connected clauses on the acceptability of sentences. For example, *Rachel disappeared once Jon had fallen asleep* is fine, but **Rachel had disappeared once Jon fell asleep* is unacceptable. A theory of acceptability is developed and its implications for interpretation discussed. Successful factoring of the linguisitic knowledge into a general, syntactic component and a lexical component dependent on the properties of individual connectives clarifies the intrepretation problem. Finally, a computer model of the theory, which serves as a workbench and confirms the theory's behavior, is demonstrated.

This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for this research has been provided by a grant from the Kapor Family Foundation and by NSF Grant DCR-85552543 under a Presidential Young Investigator's Award to Professor Robert C. Berwick. Useful guidance and commentary during this research were provided by Bob Berwick and Norbert Hornstein. This report is a revised version of a paper appearing in the Proceedings of the Society for Cognitive Science, Ann Arbor, MI, 1989.

1 Overview

This paper elucidates the linguistic representation of temporal relations among events. It does so by examining sentences that contain two clauses connected **temporal/causal connectives**, words like *once, by the time, when,* and *before.* Specifically, the data involve the effect of the tenses of the connected clauses on the acceptability of such sentences. After a theory of acceptability is developed, then its implications for interpretation are discussed. The questions of acceptability are:

1. What knowledge about the admissible combinations of tenses in connected clauses can be represented in a general way that does not depend on the particular connective? For example, the tenses in (1) seem to be a bad combination independent of the connective.

   (1) * Rachel disappears { *when, *once, *by the time, *before } Jon will fall asleep

2. What limitations do specific connectives impose on the tenses of the clauses they connect? What lexical knowledge must people have about temporal/causal connectives to identify those limitations? For example, consider (2).[1]

   (2) a. OK Rachel disappeared ONCE Jon had fallen asleep
   b. * Rachel had disappeared ONCE Jon fell asleep
   c. OK Rachel had disappeared BY THE TIME Jon fell asleep
   d. * Rachel disappeared BY THE TIME Jon had fallen asleep

To see how factoring out general constraints affects the lexical representations of connectives, consider that under a completely naive, completely lexical approach, the answer to question 2 would be: *For each temporal/causal connective, its lexical entry must represent exactly which subset of all possible combinations of tenses in two clauses the connective is compatible with.* Since English has six simple tenses (see below), a little arithmetic[2]

[1]While some speakers accept (2d), (2c) is widely preferred. In this paper sentences that are marginal but are improved by a tense shift are marked with an asterisk — more finely articulated theories of relative grammaticality must await further progress.

[2]Here is the calculation: 6 possible tenses for the matrix clause and six for the adjunct clause gives 36 possible combinations in two-clause sentences; a given temporal/causal

shows that the naive approach allows $2^{36} \approx 68$ billion possible lexical representations for connectives. This paper shows that, in fact, at most 8 different lexical representations are needed to answer question 2. This reduction is achieved by factoring general constraints that do not depend on lexical entries from specific constraints that do, i.e., from answering question 1 first. The factoring depends on a combination of results from Hornstein (in press) with new results not presented elsewhere. Hornstein's representation of tenses and the way they combine, which provides some constraint on the possible combinations, is presented in the next two sections. The remaining sections present new constraint, semantic interpretation, and a computer model.
An outline of the strategy followed in this research, which doubles as an outline of this paper, is provided below.

1. Discover and exploit as much general syntactic constraint as possible.
2. Determine which features of syntactic constructs must be represented in the lexical entries of individual connectives. (This endeavor is called **lexical syntax.**)
3. Attempt to find semantic interpretations of the lexical-syntactic features. (This endeavor is called **lexical semantics.**)
4. Implement a computer model of the theory to verify its behavior and demonstrate potential applications to natural language processing.

2 The Representation

In order to construct a formal theory explaining which tenses can be combined we need a representation of tense. The representation used here is taken from Hornstein (in press), who bases it on Comrie (1985). It is a Neo-Reichenbachian representation (Reichenbach 1947) in that its simple tense structures (STSs) relate the following three entities: the time of the event named by the verb, denoted by "E", the time of speech, denoted by "S", and a reference time, denoted by "R". The reference time R is used to

connective might be compatible with any subset of the 36 tense combinations. There are 2^{36} such subsets, and therefore at least 2^{36} possible lexical representations for a given connective.

(1) X_Y	(2) Y_X	(3) X,Y	(4) Y,X

Table 1: **The four possible relations between time points X and Y**

past	E,R R_S	(*R. disappeared*)	past perfect	E_R R_S	(*R. had disappeared*)
present	S,R R,E	(*R. disappears*)	present perfect	E_R S,R	(*R. has disappeared*)
future	S_R R,E	(*R. will disappear*)	future perfect	E_R S_R	(*R. will disappear*)

Table 2: **The six STSs that can be expressed in English verbal morphology**

locate an event with respect to another event in sentences like (2a) and (2c) above. (A mechanism for connecting tenses via the R point will be detailed below.) Each STS consists of a relation between S and R and one between R and E; S and E are not directly related. For any time points X and Y, at most one of four possible syntactic (that is, formal) relations holds between them. These are written as in Table 1.

Initially, Hornstein assumes that "_" is interpreted as temporal precedence and "," as simultaneity. Thus construed, (1) would mean *X precedes Y*, (2) would mean *Y precedes X*, and (3) and (4) would mean *X and Y are simultaneous*. Note that although Hornstein gives the same interpretation to (3) and (4), they remain syntactically distinct. For the lexical syntactic endeavor we need not commit to a specific interpretation. It is important, however, that "_" be interpreted as some partial order, call it $<$, and that "," be interpreted as some symmetric relation, call it $=$. Furthermore, $=$ must be such that $X < Y$ and $Y = Z$ together imply $X < Z$. Particular relations for $<$ and $=$ will be considered briefly in the section on lexical semantics.

There are four possible S-R relations and four possible R-E relations for a total of 16 possible simple tense structures. Of these, six[3] can be expressed in English using only verbal morphology (see Table 2). The interpretation of the past STS, for example, would be written $E = R < S$, while that of the

[3] Some of the remaining combinations occur when above the tense structures are modified by adverbs and adjuncts. See Hornstein, in press, for details.

past perfect STS would be written $E < R < S$. Under Hornstein's initial interpretation, where $<$ is precedence and $=$ is simultaneity, these STSs yield the intuitively correct ordering of events.

3 Causal/Temporal Adjunct Clauses

This section and the next correspond to stage 1 of the strategy outlined above, identifying general syntactic constraints.

The tense structure of an adjunct clause is composed with that of its matrix by **identifying** their respective S and R points. **Identifying** two points means treating them as a single entity. Consider sentence (2a). The matrix clause in (2a) is in the past tense and the adjunct clause (the one following the connective) is in the past perfect. We write the combined tense structure (CTS) for a past matrix with a past perfect adjunct as in (3),

(3)
```
E,R R_S (past matrix)
  | | |
E_R R_S (past perfect adjunct)
```

where the vertical links represent identity. If two points are to be treated as the same point they must stand in the same relationship to all other points. This becomes an issue when more than one pair of points is identified. For example, consider the CTS (4), which corresponds to sentences like (1).

(4)
```
S,R R,E  (present matrix)
|   | |
S_R R,E  (future adjunct)
```

(4) is ill-formed because the single entity formed out of S_{mat} and S_{adj} stands in two inconsistant relations to the single entity formed out of R_{mat} and R_{adj}. On the upper tier, which comes from the matrix clause's tense, we have S, R; on the lower, adjunct tier we have S_R. In this way the process of tense combination imposes a well-formedness condition on CTSs.

There is one qualification to the well-formedness condition on CTSs discussed above: in the configuration shown in (5)

(5)
```
X_Y (matrix)
| |
W,Z (adjunct)
```

the adjunct W,Z can be "harmonized" or "coerced" to W_Z. This transformation occurs only when the matrix clause is "_", the adjunct is ",", and linear order is preserved. Such a transformation accounts for the the very widespread coercibility of presents to futures. One example of such coercion is the ability of present tense clauses to be adjoined to futures, as in (6).

(6) Rachel will disappear {when, once, before} Jon falls asleep.

The combined tense structure for (6) is shown in Figure 1. Note that the adjunct clause in (6), although in the present tense, is interpreted as occurring in the future. Another example of such coercion, one not involving

```
(Recall that future = S_R R,E and present = S,R R,E)

S_R R,E (future matrix)
| | |
S_R R,E (present adjunct)
 |
COERCED
```

Figure 1: **A present tense adjunct coerced to future by a future tense matrix**

adjoined clauses this time, is the ability of present tense matrix clauses to be interpreted as futures when occurring with a future adverb, as in (7).

(7) I leave for New York tomorrow

The combination of the RS identification process for composition and the coercion discussed above permit only 16 combined tense structures. The 20 that it rules out are, like (1), clearly bad independent of the connective.[4]

4 More Constraints

Of the 16 combined tense structures permitted in Hornstein's account, seven seem to be either unattested or infelicitous. These seven are ruled out by the

[4]The identification process and the coercion process are taken from Hornstein (in press, Ch. 2), where they are discussed in slightly different terms.

two new constraints proposed below.

4.1 Interpretability Constraint

Recall that the only commitment we have made to the interpretation of "_" and "," is that the former be interpreted as some partial order that we call "<", while the latter is interpreted as some equivalence relation that we call "=". With that assumption we can sometimes use transitivity to draw conclusions about the relationship between the matrix event and the adjunct event of a CTS. For example, the CTS in (8)

(8)
```
E_R R_S (past perfect matrix)
 |  | |
E,R R_S (past adjunct)
```

yields the interpretation $E_{mat} < R = E_{adj}$ and the deduction $E_{mat} < E_{adj}$. By contrast, CTSs containing the configuration shown in (9) yield no such deduction.

(9)
```
E_R (matrix)
 |
E_R (adjunct)
```

(9) receives the interpretation $E_{mat} < R$ and $E_{adj} < R$, but it is completely uninformative about the relationship between E_{mat} and E_{adj}. The **Interpretability Constraint** says that all such CTS are infelicitous, allowing only CTSs where either $E_{adj} < E_{mat}$, $E_{mat} < E_{adj}$, or $E_{adj} = E_{mat}$ is a valid deduction. (10) shows a sentence violating the Interpretability Constraint.[5]

(10) * Rachel had disappeared before Jon had fallen asleep.

This constraint rules out 4 bad CTSs, leaving a total of 12 CTSs still possible.

[5]As usual, (10) is starred because it is not as good as a restatement with the tenses shifted to the simple past. There are certain discourse contexts in which (10) improves, but discourse considerations are beyond the scope of this paper.

4.2 Coercion Parameter & Full Interpretation

Because of the coercibility of adjunct S,R to S_R, some CTSs can be formed from more than one pair of simple tense structures. An example is shown in Figure 2. It appears that within a given language either the

```
(Recall that future = S_R R,E and present = S,R R,E)

S_R R,E (future matrix)                        S_R R,E (future matrix)
|  |  |                       is identical to  |  |  |
S_R R,E (FUTURE ADJUNCT)                       S_R R,E (PRESENT ADJUNCT)
                                                |
                                             COERCED
```

Figure 2: **Two ways of constructing the same CTS**

coerced form of the adjunct or the form that requires no coercion is consistently preferred. In English and German the coerced forms are preferred, while in Romance languages the uncoerced forms are preferred. This suggests a cross-linguistic parameterization. What's more, there seems to be an economy constraint, or perhaps some version of the **principle of full interpretation** (Chomsky 1985) guiding the choice of this parameter. In English and German using uncoerced adjuncts (i.e., future and future perfect) requires an additional auxiliary word (*will*) not required for the use of coerced adjuncts (i.e., present and present perfect). Not surprisingly, the shorter, coerced forms are preferred in English and German. (11) provides an example in English. If some version of economy or full interpretation is

(11) a. OK Rachel will disappear when Jon falls asleep *(present adj.)*
b. * Rachel will disappear when Jon will fall asleep *(future adj.)*

at work here, one would expect that languages requiring the same number of auxiliaries for the coerced and uncoerced forms would choose the coercion

	$E_{mat} < E_{adj}$	$E_{adj} < E_{mat}$	$E_{adj} = E_{mat}$
matrix	past perfect	past	past
adjunct	past	past perfect	past
matrix	present perfect	present	present
adjunct	present	present perfect	present
matrix	future perfect	future	future
adjunct	present	present perfect	present

Table 3: **Legal tense combinations arranged by interpretation**

parameter freely. This appears to be correct: Romance languages, which require auxiliaries for either present or future, use uncoerced forms.[6]

The coercion parameter rules out three of the remaining 12 CTSs, leaving only the nine acceptable ones.

5 Lexical Syntactic Features

We now turn from general constraints on all CTSs to the compatibility of particular CTSs with particular temporal/causal connectives. This constitutes stage 2 of the strategy outlined above.

Recall that the **Interpretability Constraint** permits only CTSs that yield one of the following deductions: $E_{adj} < E_{mat}$, $E_{mat} < E_{adj}$, or $E_{adj} = E_{mat}$. It turns out that **the particular deduction yielded by a given CTS is the only feature that affects its compatibility with a given temporal/causal connective.**[7] The complete body of evidence supporting this observation cannot be presented in a paper so a few illustrative examples must suffice.

Consider the connective *once*. *Once* is compatible with any CTS whose interpretation yields the deduction $E_{adj} < E_{mat}$ (see Table 3).[8] *once* is not

[6] There is some debate as to whether Chinese has a bona-fide tense system, but if it does, it too has an equal number of tense marking words either way, but chooses the opposite setting as Romance.

[7] But see the caveat on aspectual class below.

[8] *once* is also compatible with $E_{mat} = E_{adj}$, although no examples are shown.

compatible with the CTSs where $E_{mat} < E_{adj}$. Sentences (2a,b) demonstrate this. The opposite holds for *by the time*, as shown in (2c,d). *When* and *before* are compatible only with CTSs where $E_{mat} = E_{adj}$.

Since there are only three kinds of CTSs for the purposes of connective compatibility, there are at most $2^3 = 8$ possible lexical representations needed to account for connective compatibility. As noted above, this is a reduction from the $2^{36} \approx 68$ billion such representations without constraint, and $2^{16} \approx 64$ thousand based on the constraint provided by Hornstein's account of temporal adjunction.[9]

6 Lexical Semantics

This section corresponds to stage 3 of the strategy outlined above, the attempt to assign meaning to the lexical syntactic features introduced above. The following is a brief a outline of issues and possible solutions.

We have posited lexical features that assert the compatibility of a given connective with combined tense structures implying one of the three possible $E_{mat} - E_{adj}$ relations. The interpretation of these features is naturally tied to the interpretation of the $E_{mat} - E_{adj}$ relation itself. As noted above, Hornstein (in press) initially assumes that "<", the interpretation of "_", is temporal precedence and that "=", the interpretation of ",", is simultaneity. This is natural for simple tense structures, but the interpretation of = as simultaneity runs into trouble when $E_{mat} = E_{adj}$ is deduced from CTSs. To see the difficulty, first note that the connective *when* can appear only with CTSs where $E_{mat} = E_{adj}$. Then consider (12).[10] In (12a) the adjunct

(12) When they built the 39th St. Bridge...

a. ...a local architect drew up the plans
b. ...they used the very best materials
c. ...they solved most of the traffic problems

[9]The idea that only the E-E relation matters for connective compatibility was suggested in Hornstein (in press, Ch. 2), but it could not be demonstrated until the Interpretability Constraint and the Coercion Parameter were factored out.

[10]Moens and Steedman (1988) attribute (12) to Ritchie (1979).

event precedes (or overlaps) the matrix event, in (12b) they overlap, and in (12c) the matrix event precedes the adjunct event. This observation casts doubt on the possibility of interpreting the $E_{mat} = E_{adj}$ deduction, and hence the lexical features that determine connective compatibility, in any simple way. Hornstein (in press, Ch. 2) notes this and suggests that the interpretation of "," is less specific than simultaneity, allowing the connective and other contextual factors to influence it. But that approach is slightly disquieting, since the interpretation of "," as simultaneity was important for the appeal of the simple tense structures of Table 2. One way out would be to maintain simultaneity as the interpretation of ",", while loosening the notion of identification of respective S and R points. The vertical bars in a combined tense structure such as (8) might be interpreted by some relation looser than either identity or simultaneity. In that case both "," and "—" would imply =, but "," would also carry the more specific simultaneity interpretation.[11] One possible interpretation of the vertical bars (and hence of =) is that the points connected by them are temporally near enough to allow immediate causal dependency to hold between them. (See Moens & Steedman, 1988.) Another is that no relevant events intervened between E_{mat} and E_{adj}. These definitions leave the specific time frame dependent on aspectual, pragmatic, and discourse considerations, a degree of flexibility that seems to be required by the data. At the same time they preserve the natural interpretation of "," in the simple tense structures.

7 The Computer Model

This section represents stage 4 of the strategy outlined above, the development of a computer model to verify the behavior of the theory and demonstrate some potential applications.

The theory presented above was implemented as a computer program. The program operates on parse trees, building complex tense structures out of simple ones and determining whether or not they are grammatical according to the syntactic and lexical constraints of the theory. This program was linked to a simple feature-grammar parser. In addition to building the CTS for a sentence, the program lists the interpretations of the CTSs it accepts

[11]Note that the Interpretability Constraint makes sense only if $<$ is reinterpreted so that $X = Y$ implies that neither $X < Y$ and $Y < X$.

and the constraints violated by the CTSs it rejects. It's behavior on some sentences from (1) and (2) is shown below.

```
;;; * Rachel disappears when Jon will fall asleep
(compute-tense-structures
  (parse '(Rachel +s disappear when Jon will fall-asleep)))
((TS
S,R   R,E    PRESENT
| |   |
S_R   R,E    FUTURE
* CTS violates:  economy: prefer S,R   R,E    PRESENT; CDTS))

;;; OK: Rachel disappeared once Jon had fallen asleep
(compute-tense-structures
  (parse '(Rachel +ed disappear once Jon +ed have +en fall-asleep)))
((TS
E,R   R_S    PAST
  |    | |
E_R   R_S    PAST-PERFECT
interp: E(adj)<E(mat)))

;;; * Rachel had disappeared once Jon fell asleep
(compute-tense-structures
  (parse '(Rachel +ed have +en disappear once Jon +ed fall-asleep)))
((TS
E_R   R_S    PAST-PERFECT
  |    | |
E,R   R_S    PAST
* CTS violates:  connective compatibility))
```

8 Conclusions

The strategy outlined in the first section has yielded a satisfying explanation of the possible combinations of tenses and temporal/causal connectives. However, one of the stages of that strategy, finding semantic interpretations of the lexical syntactic features, remains to be fully worked out. A second direction in which this research will be pushed is toward verbal aspect, which in some cases affects the acceptability of tense combinations in interesting ways. In particular, some ideas of Hinrichs (1986) might be fruitfully applied to factor in the effect of, for example, stative verbs on the $E_{mat} - E_{adj}$ interpretation.[12] In summary, a major source of constraint on tense combinations in adjoined clauses has been uncovered and partially explored, but exciting new treasures may lie beyond the next bend.

References

Chomsky, N. (1985) *Knowledge of Language.* Praeger, New York.

Comrie, B. (1985) *Tense* Cambridge University Press, Cambridge.

Hinrichs, E. (1986) "Temporal Anaphora in Discourses of English". *Linguistics and Philosophy* 9(1): 63-82.

Hornstein, N. (forthcoming) *As Time Goes By: Tense and Universal Grammar.* MIT Press, Cambridge, MA

Moens, M. and M. Steedman (1988) "Temporal Ontology and Temporal Reference". *Computational Linguistics*, 14(2).

Reichenbach, H. (1947) reprinted 1966 in *The Elements of Symbolic Logic.* The Free Press, New York, NY.

Ritchie, G. (1979) "Temporal Clauses in English". *Theoretical Linguistics* 6:87-115.

Vendler, Z. (1967) "Verbs and Times". Ch. 4 of *Linguistics in Philosophy.* Cornell University Press, Ithaca, NY.

[12] In general, the Vendlerian classes (Vendler, 1967) of verbs may affect the acceptability of tense combinations.

CS-TR Scanning Project
Document Control Form

Date : 1 / 26 / 95

Report # Aim-1122

Each of the following should be identified by a checkmark:
Originating Department:

- [x] Artificial Intellegence Laboratory (AI)
- [ ] Laboratory for Computer Science (LCS)

Document Type:

- [ ] Technical Report (TR)
- [x] Technical Memo (TM)
- [ ] Other: ______

Document Information

Number of pages: 13 (17-images
Not to include DOD forms, printer intstructions, etc... original pages only.

Originals are:

- [x] Single-sided or
- [ ] Double-sided

Intended to be printed as :

- [ ] Single-sided or
- [x] Double-sided

Print type:

- [ ] Typewriter
- [ ] Offset Press
- [ ] Laser Print
- [ ] InkJet Printer
- [ ] Unknown
- [ ] Other: ______

Check each if included with document:

- [ ] DOD Form
- [ ] Funding Agent Form
- [ ] Cover Page
- [ ] Spine
- [ ] Printers Notes
- [ ] Photo negatives
- [ ] Other: ______

Page Data:

Blank Pages(by page number): ______

Photographs/Tonal Material (by page number): ______

Other (note description/page number):

Description :	Page Number:
IMAGE MAP	(1) UNNUMBERED TITLE PAGE
	(2-13) PAGES #'ED 1-12
	(14) SCANCONTROL
	(15-17) TRGTS

Scanning Agent Signoff:

Date Received: 1 / 26 / 95 Date Scanned: 2 / 17 / 95 Date Returned: 2 / 17 / 95

Scanning Agent Signature: Michael W. Cook

Rev 9/94 DS/LCS Document Control Form cstrform.vsd