# A Two-Level Knowledge Representation for Machine Translation: Lexical Semantics and Tense/Aspect

Bonnie J. Dorr

*Institute for Advanced Computer Studies*
*A. V. Williams Building*
*University of Maryland*
*College Park, MD 20742*
*bonnie@umiacs.umd.edu*

### Abstract

This paper proposes a two-level model that integrates tense and aspect information, based on theories by both Hornstein (in the spirit of Reichenbach) and Allen, with lexical-semantic information based on an extended version of Jackendoff's theory that includes a verb classification system proposed by Dowty and Vendler. The model is intended to be extensible to realms outside of the temporal domain (*e.g.*, the spatial domain). The integration of tense and aspect with lexical-semantics is especially critical in machine translation because of the lexical selection process during generation: there is often a number of lexical connective and tense/aspect possibilities that may be produced from a lexical semantic representation, which, as defined in the model presented here, is largely underspecified. The use of tense and aspect information allows the choice of target-language terms to be more finely tuned and the combination of event structures to be more carefully constrained.

## 1    Introduction

Recently, there has been much discussion in the literature regarding the interaction of tense and aspect with lexical-semantics (see, for example, [Bennett *et al.*, 1990], [Hinrichs, 1988], [Maybury, 1990], [Moens and Steedman, 1988], [Nakhimovsky, 1988], [Passonneau, 1988], [Pustejovsky, 1988, 1989], and [Tenny, 1989]). Among those who have studied the problem of tense and/or aspect, many have taken the work of [Reichenbach, 1947] as a starting point (see, for example, [Brent, 1988], [Moens and Steedman, 1988], [Hornstein, 1990], and [Passonneau, 1988]), while others have built on the work of [Allen, 1983, 1984] (see, for example, [Vilain *et al.*, 1990], and [Williams, 1990]), and still others have based their investigation on a combination of Allen and Reichenbach's work (see, for example, [Yip, 1985]). Among those who have studied lexical-semantic representations, many have taken the work of [Jackendoff, 1983, 1990] as a starting point (see, for example, [Brent, 1988], [Dorr, 1989, 1990a, 1990b, 1990c], [Levin and Rappaport, 1985], and [Siskind, 1989]), while others have built on the work of [Dowty, 1979] and [Vendler, 1967] (see, for example, [Bennett *et al.*, 1990], [Moens and Steedman, 1988], [Nakhimovsky, 1988], and [Passonneau, 1988]), and still others have followed [Mourelatos, 1981] and [Comrie, 1976] (see, for example, [Bach, 1986] and [Pustejovksy, 1989]).

This paper proposes a two-level model that integrates tense and aspect information, based on theories by both Hornstein (in the spirit of Reichenbach) and Allen, with lexical-semantic information based on an extended version of Jackendoff's theory that includes a verb classification system proposed by Dowty and Vendler. The model is intended to be extensible to realms outside of the temporal domain (*e.g.*, the spatial domain).
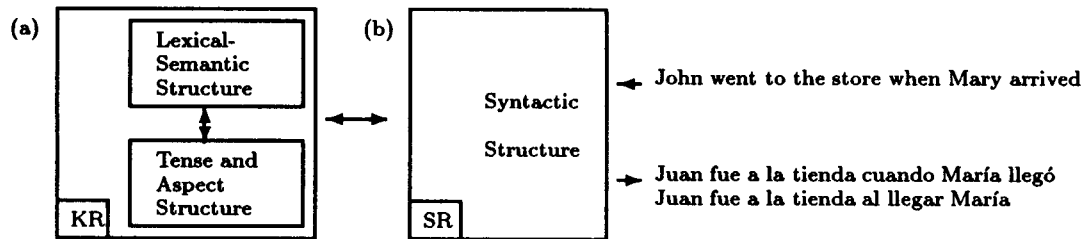
Figure 1: Two-Level Knowledge Representation of UNITRAN

The integration of tense and aspect with lexical-semantics is especially critical in machine translation because of the lexical selection process during generation: there is often a number of lexical connective and tense/aspect possibilities that may be produced from a lexical semantic representation, which, as defined in the model presented here, is largely underspecified. The use of tense and aspect information constrains the choice of target-language terms, which, in turn, limits the possibilities for the generation of tense and aspect; thus, there is a two-way communication channel between the two processes: lexical selection and tense/aspect selection.

The following section defines the dividing line between non-lexical knowledge (*i.e.*, the tense and aspect of the surface sentence) and lexical knowledge (*i.e.*, the lexical tokens that make up the surface sentence), and then discusses how these two types of knowledge are integrated in a two-level knowledge representation model for a machine translation system. Section 3 addresses the issue of cross-linguistic applicability of these two knowledge types, and section 4 discusses possible extensions of temporal knowledge to the spatial domain.

# 2 Interaction of Tense and Aspect with Lexical Semantics

The hypothesis proposed by [Tenny, 1987] is that the mapping between cognitive structure and syntactic structure is governed by aspectual properties. The implication is that lexical-semantic knowledge exists at a level that does not include aspectual or temporal information (though these two types of knowledge may depend on each other in some way). This is the view that is adopted here: it is assumed that lexical semantic knowledge consists of such notions as predicate-argument structure, well-formedness conditions on predicate-argument structures, and procedures for lexical selection of surface-sentence tokens; all other types of knowledge must be represented at some other level.

The lexical-semantic representation that is adopted as the interlingua for the UNITRAN machine translation system [Dorr, 1989, 1990a, 1990b, 1990c] is an extended version of *lexical conceptual structure* (henceforth, LCS) (see [Jackendoff, 1983, 1990]). This representation is the basis for the lexical-semantic level that is included in the knowledge representation (KR) component (see figure 1(a)). The second level that is included in this component is the tense and aspect structure.

In addition to the KR component, there is also a syntactic representation (SR) component (see figure 1(b)) that is used for manipulating the syntactic structure of a sentence. Together, the KR and SR operate bidirectionally in order to analyze the source-language sentence and synthesize the target-language sentence. We will omit the discussion of the

SR component of UNITRAN (see, for example, [Dorr, 1987]) and will concern ourselves only with the KR component for the purposes of this paper.

The translation example shown here illustrates the fact that the English sentence *John went to the store when Mary arrived* can be translated in two ways in Spanish. This example is addressed further throughout this paper. The remainder of this section defines the dividing line between non-lexical knowledge (*i.e.*, tense and aspect) and lexical knowledge (*i.e.*, properties of predicates and their arguments), and discusses how these two types of knowledge might be integrated in a two-level knowledge representation model for a machine translation system.

## 2.1 Tense and Aspect Structure

The information required for the realization of tense and aspect is considered to be outside of the scope of the lexical knowledge. Aspect is taken to have two components, one that distinguishes between states and events,[1] and the other that defines the perspective (*i.e.*, simple, progressive, and perfective). (See, for example, [Dowty, 1979] and [Vendler, 1967].) Tense, on the other hand, is taken to be the external time relationship between a given situation and others. (See, for example, [Bennett *et al.*, 1990]).

In the example of figure 1, the source- and target-language sentences consist of two event structures, each of which is associated with its own tense and aspect structure. In the case of *go* (*John went to the store*), the event is associated with the Reichenbachian Basic Tense Structure (BTS) E,R_S, which indicates that the event is in the past.[2] The aspect of this clause is "simple" (as opposed to progressive or perfective). In the case of *arrive* (*Mary arrived*), the event is associated with the same Reichenbachian temporal representation (E,R_S) and aspect (simple), since it too is in the simple past tense. As for relating these two events, the approach adopted here is based on a neo-Reichenbachian framework proposed by [Hornstein, 1990] in which the basic tense structures are organized into a complex tense structure (CTS) as follows: the first event (*i.e.*, the matrix clause) is written over the BTS of the second event (*i.e.*, the adjunct clause) and the S and R points are then associated.[3] The entire temporal/aspectual structure for this example would be specified as follows:

---

[1] We will see in section 2.2 that events are further subdivided into activities, achievements, and accomplishments.

[2] It is assumed that the reader is familiar with the Reichenbachian framework, which postulates three theoretical entities: S (the moment of speech), R (a reference point), and E (the moment of the event). The key idea is that certain linear orderings of the three time points get grammaticalized into six basic tenses in English. The corresponding Basic Tense Structures are:

| | |
|---|---|
| S,R,E | present |
| E,R_S | past |
| S_R,E | future |
| E_S,R | present perfect |
| E_R_S | past perfect |
| S_E_R | future perfect |

The S, R, and E points may be separated by a line (in which case, the leftmost point is interpreted as temporally earlier than the other) or by a comma (in which case, the points are interpreted as contemporaneous).

[3] In the general case, the association of the S and R points may force the $R_2$ point to be moved so that it is aligned with the $R_1$ point. The $E_2$ point is then placed accordingly.

$$(1) \quad \begin{bmatrix} E_1, R_1\_S_1 \\ \quad | \quad | \\ E_2, R_2\_S_2 \\ \text{aspect}_1 = \text{simple} \\ \text{aspect}_2 = \text{simple} \end{bmatrix}$$

Both tense and aspect are considered to be non-lexical in that they are determined by factors relating not to the lexical-semantic structure or particular lexical tokens of the surface sentence, but to the temporal/aspectual features of the context surrounding the event coupled with certain linguistically motivated constraints on the tense structure of the sentence. In particular, it has been persuasively argued by [Hornstein, 1990] that all sentences containing a matrix and adjunct clause are subject to a linguistic (syntactic) constraint on tense structure *regardless* of the lexical tokens included in the sentence. For example, Hornstein's linguistic Constraint on Derived Tense Structures (CDTS) requires that the association of S and R points not involve crossover in a complex tense structure:

(2) * John went to the store when Mary arrives.

$$\begin{bmatrix} E_1, R_1\_S_1 \\ \qquad \diagdown\kern-0.4em\diagdown \\ S_2, R_2, E_2 \\ \text{aspect}_1 = \text{simple} \\ \text{aspect}_2 = \text{simple} \end{bmatrix}$$

Here, the association of $R_2$ and $R_1$ violates the CDTS, thus ruling out the sentence.

Note that this linguistic constraint is a syntactic restriction on the manipulation of tense structures, not on the temporal interpretation of tensed sentences. Thus, the constraint holds regardless of the lexical token that is chosen as the connective between the two events.

$$(3) \quad \text{* John went to the store} \quad \left\{ \begin{array}{l} \text{as} \\ \text{before} \\ \text{after} \\ \text{as soon as} \\ \text{while} \end{array} \right\} \text{Mary arrives.}$$

The connecting word that relates the two events must be selected independently of the temporal/aspectual structure associated with the sentence since the order of $E_1$ and $E_2$ in a given CTS does not necessarily correspond to the order imposed by the interpretation of the connective. For example, the CTS for *John went to the store before Mary had arrived* is identical to the CTS for *John went to the store after Mary had arrived*, even though $E_1$ is placed linearly *after* $E_2$ in both cases:

$$(4) \quad \begin{bmatrix} E_1, R_1\_S_1 \\ \quad | \quad | \\ E_2\_R_2\_S_2 \\ \text{aspect}_1 = \text{simple} \\ \text{aspect}_2 = \text{perfective} \end{bmatrix}$$

Thus, it is assumed that the knowledge required to determine the temporal/aspectual structure associated with a sentence exists at a level that is independent from the lexical-semantic knowledge required to select the appropriate lexical items for the surface sentence.

The claim that there is a separation of temporal/aspectual knowledge from lexical-semantic knowledge is further strengthened if one considers that a given lexical-semantic representation may correspond to more than one tense and aspect structure, depending

on the context of the linguistic utterance. For example, suppose we are given the fact that John went to the store before Mary arrived; this fact can be represented as a single lexical-semantic structure, but it may be associated with a number of tense and aspect structures, each of which corresponds to a different surface-sentence utterance:

(5) John went to the store before Mary arrived.

$$\begin{bmatrix} E_1, R_1\_S_1 & aspect_1 = simple \\ \mid \quad \mid & \\ E_2, R_2\_S_2 & aspect_2 = simple \end{bmatrix}$$

(6) John went to the store before Mary had arrived.

$$\begin{bmatrix} E_1, R_1\_S_1 & aspect_1 = simple \\ \mid \quad \mid & \\ E_2\_R_2\_S_2 & aspect_2 = perfective \end{bmatrix}$$

(7) John had gone to the store before Mary arrived.

$$\begin{bmatrix} E_1\_R_1\_S_1 & aspect_1 = perfect \\ \mid \quad \mid & \\ E_2, R_2\_S_2 & aspect_2 = simple \end{bmatrix}$$

(8) John had gone to the store before Mary had arrived.

$$\begin{bmatrix} E_1\_R_1\_S_1 & aspect_1 = perfective \\ \mid \quad \mid & \\ E_2\_R_2\_S_2 & aspect_2 = perfective \end{bmatrix}$$

All of these surface realizations are perfectly valid given that they are consistent with our temporal knowledge, $i.e.$, that the "going to the store" event occurs before the "arriving" event. Thus, the lexical-semantic structure for these two events does not constrain the choice of potential tense and aspect structures with which it may be associated. This provides further evidence that temporal/aspectual knowledge exists at a level that is independent from lexical-semantic knowledge.

Note that Hornstein's neo-Reichenbachian theory crucially relies on an asymmetry between the matrix and adjunct clauses. Thus, there is an important distinction between [Hornstein, 1990], in which the asymmetrical property is fundamental to the theory, and [Yip, 1985], in which the asymmetrical property is entirely abandoned. I suggest that Hornstein's intuition is the correct one given that we cannot arbitrarily interchange the matrix and adjunct clauses. For example, Yip's theory predicts that we should be able to replace "$E_1$ after $E_2$" with "$E_2$ before $E_1$," which is not always the case:

(9)  (i)  John will go to the store after Mary has arrived.

    (ii)  * Mary has arrived before John will go to the store.

(10)  (i)  John will go to the store after Mary arrives.

    (ii)  * Mary arrives before John will go to the store.

Given this asymmetrical property, it would not be possible to randomly select a matrix/adjunct order and an appropriate temporal connective for a surface sentence solely on the basis of lexical information. What is needed is the temporal relation between the two events and the constraints on their combination before it is possible to derive the surface structure of the sentence. In addition, the aspectual information must be determined before the two events can be combined since this information will be necessary for retrieving the tense structure ($e.g.$, simple $vs.$ progressive) and selecting the lexical connective ($e.g.$, when $vs.$ while). We will return to the problem of the interaction between tense/aspect and lexical semantics in section 2.3, but first we will turn to a brief description of the lexical-semantic representation that is used as the interlingua for the machine translation system.

## 2.2 Lexical-Semantic Structure

Lexical-semantic structure exists at a level of knowledge representation that is distinct from that of tense and aspect in that it encodes information about predicates and their arguments, plus the potential realization possibilities in a given language. In terms of the representation proposed by [Jackendoff, 1983, 1990], the lexical-semantic structures for the two events of figure 1 would be the following:

(11) (i)  [$_{Event}$ GO$_{Loc}$ ([$_{Thing}$ John], [$_{Position}$ TO$_{Loc}$ ([$_{Thing}$ John], [$_{Location}$ Store])])]

(ii)  [$_{Event}$ GO$_{Loc}$ ([$_{Thing}$ Mary], [$_{Position}$ TO$_{Loc}$ ([$_{Thing}$ Mary], [$_{Location}$ e])])][4]

Although temporal connectives are not discussed in the theory proposed by [Jackendoff, 1983, 1990], it is assumed that these two structures are related by means of a lexical-semantic token corresponding to the temporal relation between the two events.

As it turns out, this Jackendoff-style lexical-semantic representation is sufficient for the purposes of producing an interlingual representation for machine translation. (See, for example, [Dorr, 1990b, 1990c].) However, a richer lexical-semantic representation is necessary in order to accommodate a framework in which predicates of different temporal/aspectual categories are readily distinguished. Although the lexical-semantic representation provided by Jackendoff distinguishes between events and states, this distinction alone is not sufficient for choosing among similar predicates that occur in different temporal/aspectual categories. In particular, events can be further subdivided into more specific types so that *non-atomic* events (*i.e.*, events that are allowed to have an extended interpretation) such as *destroy* can be distinguished from *atomic* events (*i.e.*, events that never have an extended interpretation) such as *obliterate*. This distinction is not captured in Jackendoff's framework, but it is a crucial distinction given that these two similar words cannot be interchangeably used in all temporal/aspectual contexts:

(12) (i)  John destroyed the house $\left\{ \begin{array}{l} \text{for an hour.} \\ \text{until Jack arrived.} \end{array} \right\}$

(ii)  * John obliterated the house $\left\{ \begin{array}{l} \text{for an hour.} \\ \text{until Jack arrived.} \end{array} \right\}$

Such distinctions are omitted in [Yip, 1985], where all events are merged under the single heading *dynamic*, and a constraint on temporal interpretations is applied uniformly across all verbs in this category. However, as we have seen above in example (12), it cannot be the case that the temporal interpretations of all dynamic verbs adhere to the same constraints given that the temporal/aspectual distributions are not identical.

A number of lexical-semantic representations have been proposed that more readily accommodate temporal/aspectual distinctions. In particular, [Dowty, 1979] and [Vendler, 1967] have proposed an aspectually oriented lexical-semantic structure that provides a four-way classification system for verbs: states, activities, achievements, and accomplishments, each of which has a different degree of telicity (*i.e.*, culminated *vs.* nonculminated), and/or atomicity (*i.e.*, point *vs.* extended).[5] A similar scheme has been suggested by

---

[4] The empty location denoted by $e$ corresponds to an unrealized argument of the predicate *arrive*.

[5] Dowty's version of this classification collapses achievements and accomplishments into a single event type called a *transition*, which covers both the point and extended versions of the event type. The rationale for this move is that all events have *some* duration, even in the case of so-called punctual events, depending on the granularity of time involved. (See [Passonneau, 1988] for an adaptation of this scheme as implemented in the PUNDIT system.) For the purposes of this discussion, we will maintain the distinction between achievements and accomplishments.

| Dowty; Vendler; Passonneau | Bennett; Moens & Steedman | Mourelatos; Comrie; Bach; Pustejovsky | Jackendoff | Examples |
|---|---|---|---|---|
| State | [-d] | State | State (BE) | be, like, know |
| Activity (point) | [+d,-t,+a] | Process | Event (GO, STAY) | tap, wink |
| Activity (extended) | [+d,-t,-a] | Process | Event (GO, STAY) | swim, run |
| Achievement | [+d,+t,+a] | Event | Event (GO, STAY) | obliterate, kill |
| Accomplishment | [+d,+t,-a] | Event | Event (GO, STAY) | destroy, give |

Figure 2: Proposals for Lexical-Semantic Frameworks that Accommodate Tense/Aspect

[Bach, 1986] and [Pustejovksy, 1989] (following [Mourelatos, 1981] and [Comrie, 1976]) in which actions are classified into states, processes, and events.

In light of these observations, the lexical-semantic structure adopted for UNITRAN is an augmented form of Jackendoff's representation in which events are distinguished from states (as before), but they are further subdivided into activities, achievements, and accomplishments. The subdivision is achieved by means of three features proposed by [Bennett et al., 1990] following the framework of [Moens and Steedman, 1988] (in the spirit of [Dowty, 1979] and [Vendler, 1967]): ±dynamic (i.e., events vs. states, as in the Jackendoff framework), ±telic (i.e., culminative events (transitions) vs. nonculminative events (activities)), and ±atomic (i.e., point events vs. extended events). This featural system is imposed on top of the lexical-semantic framework proposed by Jackendoff. For example, the primitive GO would be annotated with the features [+d,+t,-a] for the verb destroy, but [+d,+t,+a] for the verb obliterate, thus providing the appropriate distinction for cases such as (12).

Figure 2 relates the four types of lexical-semantic frameworks outlined above. Note that the system of features proposed by [Bennett et al., 1990] and [Moens and Steedman, 1988] provide the finest tuning given that five distinct categories of predicates are identified by the feature settings. (This system is essentially equivalent to the Dowty/Vendler proposal, but features are used to distinguish the categories more precisely.) In the next section, we will see how the tense and aspect structure described in section 2.1 and the lexical-semantic representation described in this section are combined to provide the framework for generating a target-language surface form.

## 2.3 Combining Tense/Aspect with Lexical Semantics

The tense/aspect component of UNITRAN operates in tandem with the lexical-semantic component in order to provide a surface form that corresponds to the interlingual representation. Tense information is linked with the lexical-semantic representation by means of Allen's temporal relations.[6] These temporal relations are assumed to be determined from the context in which the source-language sentence is uttered or, perhaps, from some knowledge source such as a database with temporal information.

For example, if we determine from a knowledge source that event $E_1$ *John went to the store* and event $E_2$ *Mary arrived* have both occurred in the past, then the time of the

---

[6] It is assumed that the reader is familiar with Allen's 13 notational relations: > (after), < (before), = (equal), m (meets), mi (is met by), o (overlaps), oi (is overlapped by), d (during), di (contains), s (starts), si (is started by), f (finishes), and fi (is finished by).

linguistic utterance S is *after* the two event times.[7] This means that the only possible BTS's (for both $E_1$ and $E_2$) are: E,R_S (past), E_S,R (present perfect), and E_R_S (past perfect). In each of these three cases, the event time E and the speech time S are separated by (at least one) line, thus providing a temporal interpretation in which E occurs before S.

Figure 3 illustrates the combination of the two BTS's into nine possible complex tense structures. (One component of the aspectual representation, simple *vs.* perfect, is included as well.) The CDTS rules out four of the nine possibilities leaving the following five cases:[8]

(13)  John went to the store when Mary arrived.

(14)  John went to the store when Mary had arrived.

(15)  John has gone to the store when Mary has arrived. [as in: Typically, John has gone ...]

(16)  John had gone to the store when Mary arrived.

(17)  John had gone to the store when Mary had arrived.

Now that the constraint proposed by Hornstein has pared down the possibilities for the tense combinations, we can further constrain the choices of surface sentences by selecting the most accurate description of the temporal relation between the two events for the connective *when*. Suppose we have the additional information from the knowledge source that the "going to the store" event occurs before the "arriving" event; in terms of Allen's notation, this would be specified as $E_1 < E_2$. The two sentences that guarantee this relation are (16) and (17). We can ensure that only these two realizations are selected for the *when* connective by using the $E_1 < E_2$ relation as an index into a table that associates the aspectual information of the two events with the *when* connective. The aspectual information in this table includes the featural specifications (*i.e.*, ±dynamic, ±telic, ±atomic) for the two lexical-semantic tokens (*i.e.*, $GO_{Loc}$ in both cases) as well as the perspective of the two events (*i.e.*, simple, progressive, or perfective). A portion of such a table is shown in figure 4. From this table we see that there are only two possible aspectual realizations for the two $GO_{Loc}$ events in the current example: since both events are associated with the features [+d,+t,-a] and since there are only five legal tense/aspect combinations to choose from, the only admissible aspectual realizations are a perfective matrix clause with a simple adjunct clause, or a perfective matrix clause with a perfective adjunct clause. Thus, sentences (16) and (17) are selected as legal possibilities for the surface sentence.

Both [Brent, 1990] and [Yip, 1985] have attempted to compile a table along the lines of the one shown in figure 4. However, in the case of Yip, the perfective aspect is omitted, thus excluding sentences such as (16) and (17) as possible surface sentence forms. In the case of Brent, the table is calculated irrespective of the lexical connective, thus giving rise to spurious temporal assignments such as $E_1 = E_2$ for the temporal interpretation of sentences such as *John went to the store before Mary arrived.* Furthermore, only punctual events are considered; interval events are entirely ignored in Brent's analysis of temporal connectives. Neither Yip nor Brent take telicity or atomicity into account in the construction of the connective table. However, as we will see in section 3, these features are important for providing a cross-linguistically applicable framework for tense and aspect in a machine translation model.

---

[7]We are assuming that the time of the linguistic utterance S refers to the present time.

[8]Analogous results would be obtained if we were to switch the matrix and adjunct clauses for this example, although this is not always the case.

i. Past / Past : John went to the store when Mary arrived.

$$\left[\begin{array}{ll} E_1, R_1\_S_1 & \text{aspect}_1 = \text{simple} \\ \quad |\quad\; | & \\ E_2, R_2\_S_2 & \text{aspect}_2 = \text{simple} \end{array}\right]$$

ii. Past / Present Perfect : * John went to the store when Mary has arrived.

$$\left[\begin{array}{ll} E_1, R_1\_S_1 & \text{aspect}_1 = \text{simple} \\ E_2\_S_2, R_2 & \text{aspect}_2 = \text{perfective} \end{array}\right]$$

iii. Past / Past Perfect : John went to the store when Mary had arrived.

$$\left[\begin{array}{ll} E_1, R_1\_S_1 & \text{aspect}_1 = \text{simple} \\ \quad |\quad\; | & \\ E_2\_R_2\_S_2 & \text{aspect}_2 = \text{perfective} \end{array}\right]$$

iv. Present Perfect / Past : * John has gone to the store when Mary arrived.

$$\left[\begin{array}{ll} E_1\_S_1, R_1 & \text{aspect}_1 = \text{perfective} \\ E_2, R_2\_S_2 & \text{aspect}_2 = \text{simple} \end{array}\right]$$

v. Present Perfect / Present Perfect : John has gone to the store when Mary has arrived.

$$\left[\begin{array}{ll} E_1\_S_1, R_1 & \text{aspect}_1 = \text{perfective} \\ \quad |\quad\; | & \\ E_2\_S_2, R_2 & \text{aspect}_2 = \text{perfective} \end{array}\right]$$

vi. Present Perfect / Past Perfect : * John has gone to the store when Mary had arrived.

$$\left[\begin{array}{ll} E_1\_S_1, R_1 & \text{aspect}_1 = \text{perfective} \\ E_2\_R_2\_S_2 & \text{aspect}_2 = \text{perfective} \end{array}\right]$$

vii. Past Perfect / Past : John had gone to the store when Mary arrived.

$$\left[\begin{array}{ll} E_1\_R_1\_S_1 & \text{aspect}_1 = \text{perfect} \\ \quad |\quad\; | & \\ E_2, R_2\_S_2 & \text{aspect}_2 = \text{simple} \end{array}\right]$$

viii. Past Perfect / Present Perfect : * John had gone to the store when Mary has arrived.

$$\left[\begin{array}{ll} E_1\_R_1\_S_1 & \text{aspect}_1 = \text{perfective} \\ E_2\_S_2, R_2 & \text{aspect}_2 = \text{perfective} \end{array}\right]$$

ix. Past Perfect / Past Perfect : John had gone to the store when Mary had arrived.

$$\left[\begin{array}{ll} E_1\_R_1\_S_1 & \text{aspect}_1 = \text{perfective} \\ \quad |\quad\; | & \\ E_2\_R_2\_S_2 & \text{aspect}_2 = \text{perfective} \end{array}\right]$$

Figure 3: Nine Possible Tense/Aspect Combinations for Two Events Occurring in the Past

| When | Matrix | | Adjunct | |
|---|---|---|---|---|
| Relation | Type | Perspective | Type | Perspective |
| $E_1 = E_2$ | [$\pm$d,-t,$\pm$a] | simple, progressive, perfective | [+d,+t,$\pm$a] | simple, progressive, perfective |
| $E_1 < E_2$ | [$\pm$d,-t,$\pm$a] | perfective | [+d,+t,$\pm$a] | simple, perfective |
| $E_1 > E_2$ | [$\pm$d,+t,$\pm$a] | simple | [+d,+t,$\pm$a] | simple |
| $E_1$ mi $E_2$ | [$\pm$d,+t,$\pm$a] | simple | [+d,+t,$\pm$a] | simple |
| $E_1$ di $E_2$ | [$\pm$d,+t,$\pm$a] | simple, progressive | [+d,+t,$\pm$a] | simple |
| $E_1$ fi $E_2$ | [+d,+t,$\pm$a] | progressive | [+d,+t,$\pm$a] | simple |
| $E_1$ si $E_2$ | [-d,-t,$\pm$a] | simple | [+d,+t,$\pm$a] | simple |
| $E_1$ m $E_2$ | [-d,$\pm$t,$\pm$a] | simple | [+d,+t,$\pm$a] | simple |

Figure 4: Temporal Relations Allowed for the *When* Connective

Now that we have looked at the constraints on the choice of temporal/aspectual realizations for the matrix and adjunct clauses with respect to the *when* temporal connective, there is still the question of how the temporal connective is selected in the first place. For example, without any additional information, we are unable to determine, unambiguously, the relation between the clauses of the sentences shown in figure 1:

$$
(18) \quad \left[ \begin{array}{l} [_{\text{Event}}\ \text{GO}_{\text{Loc}}\ ([_{\text{Thing}}\ \text{John}],\ [_{\text{Position}}\ \text{TO}_{\text{Loc}}\ ([_{\text{Thing}}\ \text{John}],\ [_{\text{Location}}\ \text{Store}])])] \\ [_{\text{Event}}\ \text{GO}_{\text{Loc}}\ ([_{\text{Thing}}\ \text{Mary}],\ [_{\text{Position}}\ \text{TO}_{\text{Loc}}\ ([_{\text{Thing}}\ \text{Mary}],\ [_{\text{Location}}\ e])])] \\ \text{Potential Temporal Relations}: >, =, \text{mi}, \end{array} \right]
$$

In particular, the connective *when* may correspond to any of the three relations (*i.e.*, $>$, $=$, or mi) depending on the intended temporal interpretation of the associated events.

The mapping between temporal relations and temporal connectives is a problem in both directions: on the one hand, connectives that are in a one-to-many relation with Allen's temporal relations (*e.g.*, *when*, which corresponds to $>$, $=$, and mi) create a problem for the analysis of the source-language sentence; on the other hand, the temporal relations that are in a one-to-many relation with potential temporal connectives (*e.g.*, the $>$, which corresponds to *after* and *when*) create a problem for the generation of the target-language sentence. The next section addresses how to control this abundance of selection possibilities.

# 3  Classification of Connectives: Cross-Linguistic Applicability

The *when* construction has been studied extensively in the literature. In particular, it has been noticed that the aspectual properties of the clauses conjoined by *when* often change the temporal meaning of the entire sentence. (See [Moens and Steedman, 1988] and [Yip, 1985] among others.) Allen's temporal logic falls short in this regard: it does not capture well-known patterns of tense implications based on aspectual distinctions.[9]

Some examples of *when* constructions (taken from [Yip, 1985]) are shown here:[10] (Note that the word *when* potentially maps to more than one temporal relation if we adopt Allen's framework.)

(19)  Simple process with simple event: $>$, $=$, mi
      John left when Mary arrived.

(20)  Progressive process with simple event: di, fi
      John was leaving when Mary arrived.

(21)  Simple state with simple event: mi, si, di, m, $>$, $<$, $=$
      John was angry when Mary arrived.

(22)  Progressive process with progressive event: $=$, f, fi, s, si, d, di, o, oi
      John was leaving when Mary was arriving.

---

[9] An example of a tense implication that is not captured in Allen's approach is that a progressive form of a process entails the negation of the perfect form (*e.g.*, *John is building a house* implies *John has not built the house*). See [Yip, 1985] for a description of other tense implications that are not captured in Allen's approach.

[10] This is a modified version of the enumerated constructions in [Yip, 1985]. Certain changes have been made to accommodate differences in interpretative judgments (by native English speakers) for the data given in Yip's presentation.

As it turns out, the temporal relations enumerated in figure 4 for the *when* connective cover the cases shown here, plus many others that are not addressed by Yip. There are two important advantages to using the featural scheme of figure 4 over the less specific scheme of [Yip, 1985]: (1) it provides a more precise specification of states and events; and (2) it includes temporal/aspectual information that is important for the realization of other types of surface-structure constituents (besides temporal connectives) such as the number of a verbal object. We will return to this second point shortly.

Although the *when* construction has been studied extensively, it has not been examined in the context of machine translation. In particular, the question of whether the tense/aspect theories of Allen, Bennett, *et al.*, Dowty, Reichenbach, Vendler, and others can be combined to provide a cross-linguistic account for the selection of tense/aspect and temporal connectives such as *when* (*e.g.*, in an interlingual translation model) has not been addressed.

Machine translation provides an appropriate testbed for trying out such theories. The problem of lexical selection during generation of the target language is the most crucial issue in this regard. For example, one must choose between the lexical tokens *cuando* and *al* when generating an equivalent Spanish temporal connective for the following two English sentences:

(23) (i)    John went to the store when Mary arrived.

(ii)   John had gone to the store when Mary arrived.

In the case of (23)(i), there are two possible translations, one that uses the connective *cuando*, and one that uses the connective *al*:

(24) (i)    Juan fue a la tienda cuando María llegó.

(ii)   Juan fue a la tienda al llegar María.

Either one of these sentences is an acceptable translation for (23)(i). However, the same is not true of (23)(ii):[11]

(25) (i)    Juan había ido a la tienda cuando María llegó.

(ii)   Juan había ido a la tienda al llegar María.

Sentence (25)(i) is an acceptable translation of (23)(ii), but (25)(ii) does not mean the same thing as (23)(ii). This second sentence implies that John has already gone to the store and come back, which is not the preferred reading.

Currently research is under way to enumerate all of the English temporal connectives (taken from Webster's on-line dictionary) in order to establish an association between these connectives and the aspectual interpretation for the matrix and adjunct events (in the manner shown in figure 4). These tables vary from language to language, but the procedure for choosing temporal connectives applies cross-linguistically once the tables for each language are compiled. For example, the table for the Spanish connective *al* would be similar to the table for the English connective *when* in figure 4 except that the specification for the "<" relation would require the matrix event to have the +telic feature (*i.e.*, the matrix action must reach a culmination). Thus, the full type entry under the matrix clause for the word *al* would be [±d,+t,±a]. This would account for the distinction between *cuando* and *al* in sentences (25)(i) and (25)(ii) above.

Space limitations do not permit the enumeration of the other temporal connective tables. Some examples of connectives that are currently being compiled into tables are: *after, as soon as, at the moment that, before, between, during, since, so long as, until,*

---

[11] I am indebted to Jorge Lobo for pointing this out to me.

*while, etc.* It is intended that these tables are to be used both for the selection of temporal connectives during the generation process (for which the relevant index into the tables would be the temporal relation and the lexical-semantic types encoded in the interlingua) and for temporal interpretation during the analysis process (for which the relevant index into the tables would be the lexical-semantic types and aspectual perspectives associated with the source-language sentence). The selection of a temporal connective, then, is simply a table look-up procedure based on the type of the events and the temporal interpretation that holds between the events. For example, if we had a [+d,-t,-a] event $E_1$ (*e.g.*, *run*) and a [+d,+t,+a] event $E_2$ (*e.g.*, *arrive*), and if we knew that the temporal relation between $E_1$ and $E_2$ was m, then searching the *when* table would fail, but searching the *until* table would succeed, thus allowing a sentence such as *John ran until Mary arrived* to be generated.

As mentioned earlier, the featural system outlined above provides a framework that is appropriate not only for the realization of temporal connectives, but also for the realization of other types of temporal/aspectual information. For example, the sentence *I stabbed Mary* could be realized in at least two ways in Spanish:

(26) (i)    Juan le dio puñaladas a Marìa

(ii)   Juan le dio una puñalada a Marìa

Both of these sentences translate literally to "John gave stab wound(s) to Mary." However, the first sentence is the repetitive version of the action (*i.e.*, there were multiple stab wounds), whereas the second sentence is the non-repetitive version of the action (*i.e.*, there was only one stab wound). This distinction is characterized by means of the atomicity feature. In (26)(i), the event is associated with the features [+d,+t,-a], whereas, in (26)(ii) the event is associated with the features [+d,+t,+a]. According to [Bennett *et al.*, 1990] (in the spirit of [Moens and Steedman, 1988]), predicates are allowed to undergo an atomicity "coercion" in which an inherently non-atomic predicate (such as *dio*) may become atomic under certain conditions. These conditions are language-specific in nature, *i.e.*, they depend on the lexical-semantic structure of the predicate in question. Given the featural scheme that is imposed on top of the lexical-semantic framework, it is easy to specify coercion functions for each language. For example, the atomicity function for the stab example would specify that a singular NP verbal object maps a [+d,-a] predicate into a [+d,+a] predicate *i.e.*, a non-atomic event becomes atomic if it is associated with a singular NP object. Thus, the notion of feature-based coercion is cross-linguistically applicable, providing a useful foundation for a model of interlingual machine translation.

# 4    Extension of the Temporal/Aspectual Framework to the Spatial Domain

In addition to investigating the cross-linguistic applicability of the temporal/aspectual framework in the context of machine translation, current research is under way to extend the representation to the spatial domain. This possibility has also been investigated by [Mukerjee and Joe, 1990], in which the interval logic model of [Allen, 1983] has been extended to be applicable to the spatial domain. For example, the one-dimensional interval relation C(++)B specifies that the spatial interval C is, in some sense, "after" the spatial interval B. This relation is analogous to Allen's temporal relation $C > B$, which specifies that the temporal interval C occurs after the temporal interval B. Other spatial relations

261

that are currently under investigation are: *above, before, behind, between, beyond, down, following, next to, off, on, to, under, within, etc.* It is expected that the principles that govern the relation of temporal primitives to lexical items will hold for analogous primitives in the spatial field; experiments are currently being conducted to test this hypothesis.

# 5 Summary

This paper has examined a two-level knowledge representation model for machine translation that integrates the tense and aspect information based on theories by both Hornstein (in the spirit of Reichenbach) and Allen with lexical-semantic information based on theories by Jackendoff in the spirit of Dowty and Vendler. We have examined the question of cross-linguistic applicability showing that the integration of tense and aspect with lexical-semantics is especially critical in machine translation when there are a number of temporal/aspectual possibilities that may be generated from a lexical semantic representation. Finally, we have discussed the possibility of extending the temporal notation to the spatial domain.

# 6 Acknowledgements

# References

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

[Allen, 1984] James F. Allen. Towards a general theory of action and time. *Artificial Intelligence*, 23(2):123–160, 1984.

[Bach, 1986] Emmon Bach. The algebra of events. *Linguistics and Philosophy*, 9:5–16, 1986.

[Bennett *et al.*, 1990] Winfield S. Bennett, Tanya Herlick, Katherine Hoyt, Joseph Liro, and Ana Santisteban. A computational model of aspect and verb semantics. *Machine Translation*, 4(4):247–280, 1990.

[Brent, 1988] Michael R. Brent. Decompositional semantics and argument expression in natural language. Master's thesis, Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, Cambridge, MA, 1988.

[Brent, 1990] Michael R. Brent. A simplified theory of tense representations and constraints on their composition. In *Proceedings of the 28th Annual Conference of the Association for Computational Linguistics*, University of Pittsburgh, Pittsburgh, PA, 1990.

[Comrie, 1976] Bernard Comrie. *Aspect*. Cambridge University Press, Cambridge, England, 1976.

[Dorr, 1987] Bonnie J. Dorr. Unitran: A principle-based approach to machine translation. Master's thesis, MIT AI Technical Report 1000, Department of Electrical Engineering and Computer Science, Cambridge, MA, 1987.

[Dorr, 1989] Bonnie J. Dorr. Lexical conceptual structure and generation in machine translation. In *MIT AI Memo 1160, Proceedings of the Ninth Annual Conference of the Cognitive Science Society*, Ann Arbor, MI, 1989.

[Dorr, 1990a] Bonnie J. Dorr. Solving thematic divergences in machine translation. In *Proceedings of the 28th Annual Conference of the Association for Computational Linguistics*, pages 127–134, University of Pittsburgh, Pittsburgh, PA, 1990.

[Dorr, 1990b] Bonnie J. Dorr. A cross-linguistic approach to machine translation. In *Proceedings of the Third International Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages*, pages 13–32, Linguistics Research Center, The University of Texas, Austin, TX, 1990.

[Dorr, 1990c] Bonnie J. Dorr. *Lexical Conceptual Structure and Machine Translation*. PhD thesis, Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, Cambridge, MA, 1990.

[Dowty, 1979] David Dowty. *Word Meaning and Montague Grammar*. Reidel, Dordrecht, Netherlands, 1979.

[Hinrichs, 1988] Erhard W. Hinrichs. Tense, quantifiers, and contexts. *Computational Linguistics*, 14(2):3–14, 1988.

[Hornstein, 1990] Norbert Hornstein. *As Time Goes By*. MIT Press, Cambridge, MA, 1990.

[Jackendoff, 1983] Ray S. Jackendoff. *Semantics and Cognition*. MIT Press, Cambridge, MA, 1983.

[Jackendoff, 1990] Ray S. Jackendoff. *Semantic Structures*. MIT Press, Cambridge, MA, 1990.

[Levin and Rappaport, 1985] Beth Levin and Malka Rappaport. The formation of adjectival passives. Lexicon Project Working Papers 2, Massachusetts Institute of Technology, Center for Cognitive Science, Cambridge, MA, 1985.

[Maybury, 1990] Mark T. Maybury. Using discourse focus, temporal focus, and spatial focus to plan narrative text. In *Proceedings of the Fifth International Workshop on Natural Language Generation*, pages 70–78, Dawson, Pennsylvania, 1990.

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

[Mourelatos, 1981] Alexander Mourelatos. Events, processes and states. In *Tense and Aspect*, Academic Press, New York, NY, 1981.

[Mukerjee and Joe, 1990] Amitabha Mukerjee and Gene Joe. A qualitative model for space. In *Proceedings of the Ninth Annual Conference of the American Association of Artificial Intelligence*, Boston, MA, 1990.

[Nakhimovsky, 1988] Alexander Nakhimovsky. Aspect, aspectual class, and the temporal structure of narrative. *Computational Linguistics*, 14(2):29–43, 1988.

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

[Pustejovksy, 1988] James Pustejovksy. The geometry of events. in Lexicon Project Working Papers 24, Massachusetts Institute of Technology, Center for Cognitive Science, Cambridge, MA, 1988.

[Pustejovksy, 1989] James Pustejovksy. The semantic representation of lexical knowledge. In *Proceedings of the First Annual Workshop on Lexical Acquisition, IJCAI-89*, Detroit, Michigan, 1989.

[Reichenbach, 1947] H. Reichenbach. *Elements of Symbolic Logic*. Macmillan, London, 1947.

[Siskind, 1989] Jeffrey Mark Siskind. Decomposition. MIT Computer Science Area Exam Paper, Massachusetts Institute of Technology, Cambridge, MA, 1989.

[Tenny, 1987] Carol Tenny. *Grammaticalizing Aspect and Affectedness*. PhD thesis, Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, Cambridge, MA, 1987.

[Tenny, 1989] Carol Tenny. The aspectual interface hypothesis. Lexicon Project Working Papers 31, Massachusetts Institute of Technology, Center for Cognitive Science, Cambridge, MA, 1989.

[Vendler, 1967] Zeno Vendler. Verbs and times. *Linguistics in Philosophy*, pages 97–121, 1967.

[Vilain et al., 1990] Marc Vilain, Henry Kautz, and Peter van Beek. Constraint propagation algorithms for temporal reasoning: A revised report. In *Readings in Qualitative Reasoning about Physical Systems*, Morgan Kaufmann, San Mateo, CA, 1990.

[Williams, 1990] Brian C. Williams. Doing time: Putting qualitative reasoning on firmer ground. In *Readings in Qualitative Reasoning about Physical Systems*, Morgan Kaufmann, San Mateo, CA, 1990.

[Yip, 1985] Kenneth M. Yip. Tense, aspect and the cognitive representation of time. In *Proceedings of the 23rd Annual Conference of the Association for Computational Linguistics*, pages 18–26, Chicago, IL, 1985.