Title
Processing Aspectual Sematics

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 10(0)

Authors
Nirenburg, Sergei
Pustejovsky, James

Publication Date
1988

Peer reviewed

Processing Aspectual Semantics

Sergei Nirenburg

Center for Machine Translation
Carnegie-Mellon University

James Pustejovsky

Computer Science Department
Brandeis University

Abstract

A computational treatment of aspect in English is presented. A set of aspectual values is introduced and discussed. The lexical and contextual clues for determining aspectual values are determined. The structure of the entry in the main dictionary supporting aspectual (as well as other types of) analysis is illustrated. A computational framework for an aspectual analyzer is described, in which the latter is conceived as one of a group of specialist analysis modules working together, in a distributed (blackboard-oriented) computational environment.

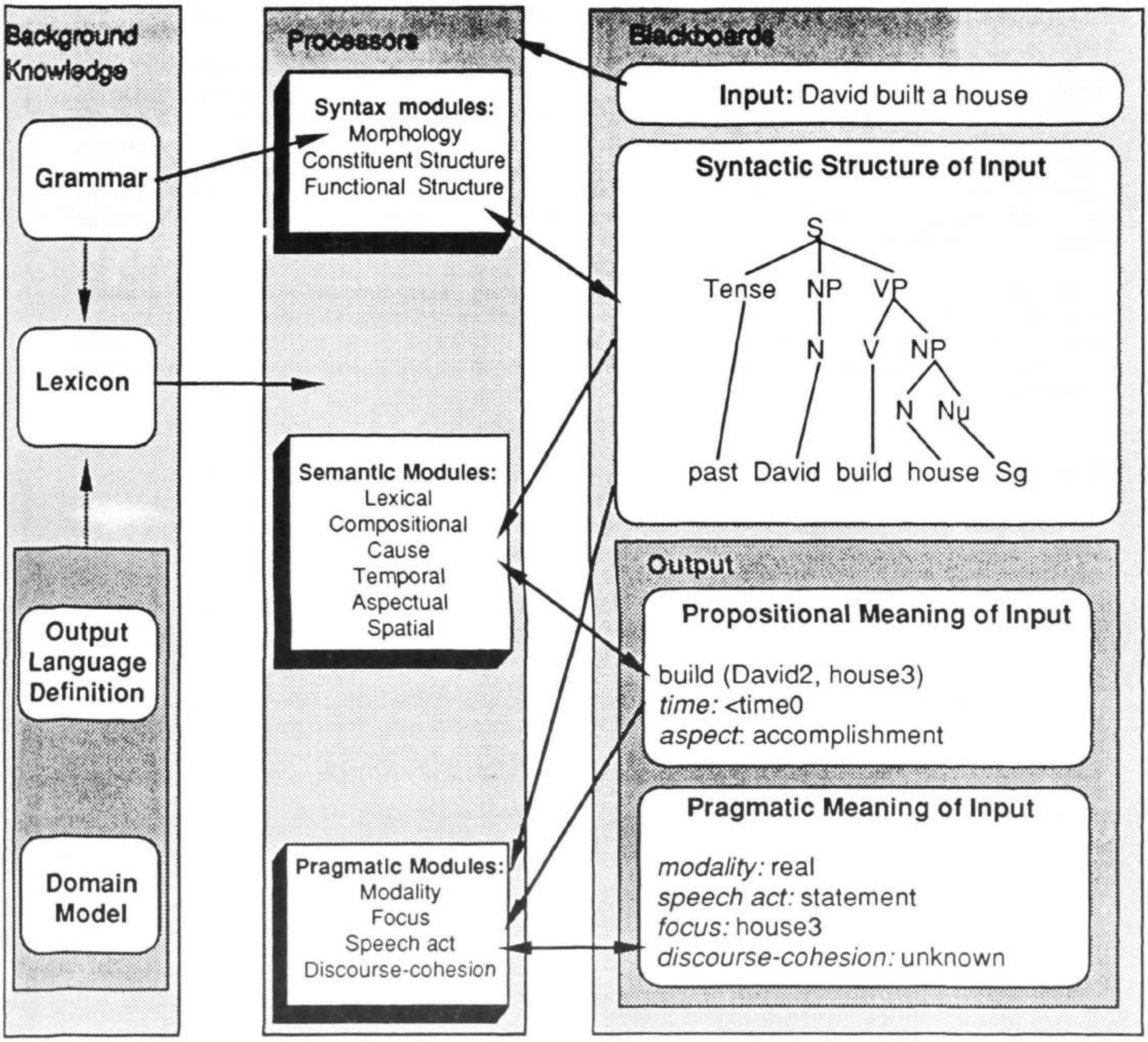
1. The Concept of Microtheories

A computational model of language behavior must provide treatment of a large number of syntactic, semantic and pragmatic phenomena. It may be realized in a set of computer programs that obtain natural language inputs, extract their meanings and represent them in a well-defined notation, after which they react accordingly to the message in the input. Of course, some types of reactions may have nothing to do with natural language (for instance, a robot might perform a motoric operation after having understood a verbal command). However, a number of reactions (as in dialog systems or the various text processing systems, such as, for instance, those of machine translation) involves generating natural language texts based on the extracted meanings. Thus, a complete model of language behavior must deal with recognition, representation and synthesis of natural language texts.

Significant progress has been made recently in the field with respect to the theories of syntax. Semantic and pragmatic phenomena have traditionally been less amenable to computational analysis. It does not seem plausible that an integrated semantic theory that covers all of lexical and compositional phenomena as well as the various pragmatic considerations is formulated in the near future. This assessment becomes even more evident if one recognizes the necessity of providing heuristics for automatic recognition of the multiple meaning facets of natural language texts as a part of the theory. At the same time, linguistics has accumulated a significant body of knowledge about the various semantically laden phenomena in the natural languages (cf. Raskin, 1987 for a discussion of how this body of knowledge can be applied to computational analysis).

The above suggests that one of the more feasible ways toward building a comprehensive computational model of language understanding and generation behavior in humans is to develop a large number of *microtheories* that deal with a particular linguistic phenomenon in a particular language or group of languages and then provide a computational architecture that allows the integration of the operation of all the modules based on these microtheories. Thus, one can envisage a microtheory of time, modality, speech act, causality, etc. This paper is devoted to a microtheory of *aspectual meanings.*

To integrate the microtheories we suggest the use of a version of the blackboard computational architecture, in which a number of processes co-exist and, using a variety of background knowledge modules, collectively produce a desired output. The structure of the language comprehension component of a language behavior model is illustrated in Figure 1. In this figure the processors are computational realizations of the various microtheories derived for the corresponding linguistic phenomena. These processors operate using the data from the background knowledge repositories, such as grammars and dictionaries, as well as the intermediate results stored on the universally accessible set of blackboards. A more detailed description of the model and its components see in Nirenburg and Raskin, 1987a and Nirenburg, 1987.



Background data, knowledge acquisition time —— Data, processing time

Figure 1-1: Comprehension Component of a Computational Model of Language Processing

2. Treating Aspectual Meanings: the Task

Aspectual meaning is a component of the overall meaning of a natural language utterance. It is illustrated in examples (1) -- (3) below, in which the only difference between the verb meanings is aspectual: +protracted, -culminative in (1); +protracted, +culminative in (2) and, surprisingly, -protracted, -culminative in (3).

(1) I walked in the park for an hour yesterday
(2) Yesterday I walked to work
(3) At 8 o'clock yesterday morning I was walking to work

Building a semantic analyzer that determines aspectual meanings of input utterances involves specifying a) a set of values for aspectual meanings; b) a set of rules for assigning particular values of aspect; c) the knowledge (the clues) necessary for the formulation of the conditions in these rules (to be found in a) the entries for verbs in the analysis lexicon; b) the syntactic structure of the input utterance; and c) the presence of certain aspectually significant lexical units, such as temporal modifiers or semi-auxiliary 'aspectual' verbs in the input text); and d) an architecture for the analyzer that will allow the results of one processing module (for instance, the syntactic parser) to serve as a decision aid for another module (such as the aspect analyzer).

In what follows we 1) suggest a set of aspectual meaning values, inspired by a theory of aspect described in Pustejovsky (submitted); 2) describe the architecture of a comprehensive semantic analyzer of which an aspect analyzer is a component; 3) describe the structure of the analysis lexicon; 4) survey the language material that is the basis for formulating decision rules; and 5) present a set of analysis rules for aspect.

3. A Language for Aspect-Related Analyzer Output: A Set of Aspect Values

The logical place to attach the information about aspect in a semantic representation is at the propositional level. Propositions are represented as frames that are essentially instantiations of event-types, with their arguments constrained in accordance with evidence in the input text. The arguments in a proposition representation include conceptual case roles, preconditions, effects, temporal, spatial and aspectual values. Within the current architecture, a separate analysis component is assigned the task of determining the contents of each of the above slots.

It is important to understand that our model uses predetermined calculus that determines what different types of events there can be. That is, just as a grammar defines the well-formed syntactic structures in a language, the calculus of aspect contributes to defining what a well-formed semantic structure is. Furthermore, all event-types are built recursively from two primitive event-types, *state* and *process* (see Pustejovsky 1987 for details).

The set of values defined for the aspectual slot in our system is as follows:

A **state** is an event-type which is nondecomposable and does not in itself refer to any initial or endpoint. Note that additional textual information can refer to initial or endpoints of a state or its duration. States for which such additional information is introduced will be called **bounded states**.

A **process** is a possibly ordered set of event-types each of which can be a state or a process itself.

If we consider a process where the initial and/or the final event-types are distinguished so that that single state is on a par in importance with the complement set, taken as a whole, then the resulting event is different. This phenomenon is called *headedness*, and the marked event-types, correspondingly, *heads*.

If the final element in a process is its *head*, and the process itself is semantically well-specified, then this event-type is called an **accomplishment**. Thus, in an accomplishment there is an embedded process as well as a distinguished final event or state, and both event constituents are of equal status. For example, in *John built a house* the final state of there being a house is as significant as the building process.

Analogous to the above definition, there are some processes in which the final event-type can be distinguished, but the complement set is not semantically specified through the meaning of a lexical item. Such an event-type is called **achievement.** In an achievement the process and its head do not, thus, have equal status, the head being more important.

The initial event of a process can also be distinguished as a head. These are left-headed structures. Left-headed structures arise only when the initial event can be distinguished as causing the rest of the event-types in the ensuing process. The left-headed structures are **causative-processes**.

There are event-types that allow two heads, the left ones and the right ones. If an event-type is double-headed, then the event-types that are neither initial nor final are typically underspecified semantically. The double-headed structures are called **transitions**.

4. The Aspect Analyzer

4.1. An Architecture for Distributed Analysis

The set of routines for determining aspectual values of utterances forms a functional component of a comprehensive natural language analyzer (see Section 1 above). We develop this analyzer, DIANA (Nirenburg, in preparation), as a distributed system of specialist modules working together to produce a complex output structure which is represented in our approach as a set of frames corresponding to the levels of text, sentence, clause, proposition and propositional argument denotations (cf. Nirenburg et al. 1987). Our analyzer shares these properties with the generation system DIOGENES being developed at CMU (Nirenburg, 1987). Together the two systems will form the basis of a knowledge-based machine translation system.

4.2. Knowledge Sources

Figure 1 illustrates the (static and dynamic) knowledge sources necessary to support aspectual analysis in the framework of a distributed, blackboard-based analyzer.

The lexicon, especially the semantico-pragmatic portion of its entries, is a most important knowledge source for determining aspectual values. The structure of our lexicon entries can be described as follows[1].

[1]This is an abridged version of the lexicon entry structure. See Nirenburg and Raskin, 1987b for a detailed discussion

```
AL-entry            ::= ( <SL-pattern> <meaning-pattern>*)
<SL-pattern>        ::= ( SL-Lexical-Unit <lex-info>)
<lex-info>          ::= ( (<syntactic-info>)
                          (morph <inflection-type>))
<syntactic-info>    ::= {the contents of a syntactic dictionary}
<inflection-type>   ::= {an indication of irregularities in forming
                         word forms, e.g., goose - pl. geese}
<meaning-pattern>   ::= ((token-of (value <domain-concept>))
                          (<property> (value <value>*))) |
                          (<property> (value <value>*))*
<domain-concept>    ::= {any concept in the domain model defining a sublanguage}
<property>          ::= {any relation or attribute from the domain model}
<value>             ::= {any concept or attribute (scale)
                         value in the domain model}
```

The lexicon has four kinds of meaning patterns: a) instantiations of concepts in the domain model, e.g., *computer*; b) instantiations of attributes of concepts in the domain model, e.g. *fast*; c) clues for determining the attributes of the properties comprising both the propositional and the pragmatic meaning, e.g., *finish*; and d) clues for making other semantic and pragmatic decisions, such as reference, e.g. *the*.

Group c) includes a class of verbs that have a special significance for aspectual analysis, the so-called 'phase' or 'aspectual' verbs. These verbs have a lexical meaning but no independent ontological meaning. The aspectually relevant meanings that they introduce into the overall meaning of the input can be classified as follows:

- inchoative: *start, begin, resume*
- continuative: *continue, keep*
- completive: *end, finish, complete*
- abortive: *stop, cease*
- iterative: *repeat*

4.3. Material

In this section we analyze a number of characteristic examples with respect to their aspectual values (AVs).

4.3.1. States

(4) John loves Mary
(5) John has loved Mary for 2 years

The aspectual analysis produces: *AV: state; Time: present* for (4) and *AV: bounded state; Time: begin: (NOW - 2 years)* for (5). In (4) the verb *love* is lexically specified as a state[2]. The temporal reference in which the verb is grounded will act to bound this state in any number of ways. In this case, since the tense is the present, there is no delimitation on the state. In (5) however, the present perfect together with the durative adverbial acts to left-bound the state denoted by the proposition.

4.3.2. Processes

(6) John walked yesterday
(7) John walked to work yesterday
(8) John is walking to work

(6) is analyzed as *AV: process; Time: past: yesterday*, (7) as *AV: accomplishment; Time: past:*

[2] that is, listed in the dictionary as being a token of a class of concepts which are descendants of *state* in the hierarchy of concepts that embodies the domain model

yesterday, and (8) as *AV: state S; S is a member of P in accomplishment; Time: present.*

Notice the effect of prepositional modification and that of the progressive on the aspectual value of the sentence. The prepositional phrase in (7) indicates the goal of a process with no intrinsic culmination. This goal acts to the terminate the event and shifts the aspectual type to an accomplishment. In (8), on the other hand, the progressive influences the nature of the resulting aspect-type more than the presence of the prepositional phrase. The result is a a state (see Pustejovsky (1987) for further discussion). The important thing to realize in the interpretation of this sentence is that the goal state is not entailed when the progressive applies to an accomplishment.

4.3.3. Achievements

(9) Bill won a race
(10) Bill is winning the race

(9) obtains the aspectual value of *achievement.* The aspectual class of achievement verbs is probably the most consistent since it seems to resist the modification that leads to aspect-type shifting. Thus, when the verb recognized is lexically specified as achievement, the resulting type will be the same. The one exception to this is the progressive, which has a similar effect as that mentioned above for 'walking to work'. The resulting aspect type in (10) is *state.*

4.3.4. Accomplishments

(11) Fred built a house
(12) Fred built houses for 5 years
(13) Fred was building a house

The analysis for the above brings: *AV: accomplishment* for (11); *AV: process; iterative: element of iteration:* **build a house; AV: accomplishment** for (12); and *AV: state S; Time: past* for (13).

Any accomplishment usually entails a culmination, but in (12) the aspect type is a process. This is a result of the bare plural object, which iterates over the lexical accomplishment to produce a process. Thus, a durative adverbial is permitted. When this sentence is put in the progressive form, it is stative.

4.3.5. Causative-Processes

(14) Max sent a package
(15) Max sent a package to Leo

(14) is analyzed as *AV: causative-process*, (15) as *AV: accomplishment.*

Here the process is directed by a single initiating event (the projecting), and thereafter there is no agency involved. Hence, in (15) it is not entailed that Leo *receives* the package (cf. Pustejovsky (1987)).

4.3.6. Transition

(16) John gave the book to Mary
(17) John gave books to Mary

The analyses: *AV: transition* for (16) and *AV: process, iterative; element: transition* for (17).

The most noticeable thing about these examples is the absence of any specified mode of transition. That is, what is lexically specified in the semantics of *give* is simply the beginning and end states of an event. Notice that in (17) the bare plural object shifts the interpretation to a process, as with the accomplishments.

4.4. The Analysis Rules for Aspect

The above linguistic material shows that many types of knowledge come into play in determining the aspectual value of a clause: lexical-semantic, syntactic and contextual. Specifically, this information is drawn from

- entries for verbs in the analysis lexicon
- verb tense values from the results of the syntactic module of the analyzer
- the presence of particular modifiers (adverbs and prepositional phrases) in the input (the syntactic module being responsible for determining what is modified by what)
- the presence of aspectual verbs in the input in syntactically relevant positions
- the presence, the meaning and the syntactic form of case-role holders for particular verbs (the above are determined through the operation of the syntactic and the propositional-semantic modules of the analyzer)

Specific rules and heuristics that the analysis system uses to identify the aspectual type for an input sentence include the following[3]:

- If the main predicate in an input sentence has the aspectual marker *achievement* in the dictionary, assign the aspect value *achievement*
- If the main predicate is marked in the dictionary as *state* **or** if the morphologo-syntactic analysis determines that it is in a progressive form, assign the aspect value *state*
- If the main predicate is marked as *accomplishment* **and** the direct object in the sentence is definite **or** the main predicate is marked as *process* **or** *directed-process* **and** is modified by a prepositional phrase, assign the aspect value *accomplishment*
- If the main predicate is marked as *process* **or** *directed-process* **or** a direct object is present identified as *bare-plural* **or** *mass-noun*, assign the aspect value *process*.
- If an aspectual verb is present in the sentence follow these rules:
-
    - Inchoative + (State **or** Process **or** accomplishment) = achievement; other combinations are impossible for inchoative
    - Continuative applies only to a process to yield a process
    - Completive and abortive apply to a process to yield an achievement
    - Iterative applies to achievements, accomplishments and transitions to yield processes.

5. Discussion: Future Work and Limitations

Passonneau (1987) and Moens and Steedman (1987) are two of the recent publications devoted to a similar topic.

The problem specifications in this model and Passonneau (1987) are very compatible. Still, the representation of the aspect calculus assumed here is richer than Mourelatos' (1981) typology, which Passonneau assumes, in that it proposes a larger number of primitive event-types, which, we believe provide a more accurate coverage for the aspectual phenomena. Space does not permit us to adequately motivate all our categories, but see Pustejovsky (1987) for discussion. Similarly, the set of analysis clues we employ is broader, including the input from lexical sources beyond the verb itself. Differences in

---

[3]We decided to present these heuristics not in the format used by our program but rather in a human-readable form

analyzer architecture are also quite significant.

Moens and Steedman (1987) pursue a somewhat different goal. They emphasize how aspectual processes fit into a larger temporal reasoning component to provide a richer tense system. This is also explored in Pustejovsky and Herman (1988). But our concerns in this paper are primarily elsewhere. We have addressed the problem of the lexical specification of a verbal element and how this value changes in the context of other propositional information.

The aspectual analyzer, as described above, has been prototyped. It can be tested in earnest only when at least several other modules of DIANA are implemented. As to the enhancements to the aspect module itself, we plan to extend the coverage of the aspect determination heuristics through empirical studies over large text corpora.

Bibliography

Corkill, D.D., K.Q. Gallagher, and K.E. Murray. 1986. GBB: A Generic Blackboard Development System. AAAI-86. 1008-1014.

Moens M. and M. Steedman. 1987. Temporal Ontology in Natural Language. ACL-87.

Mourelatos, A.P.P. 1981. Events, Processes, and States. In: P. Tedeschi and A. Zaenen. Syntax and Semantics, vol. 14. Tense And Aspect. NY: Academic Press.

Nirenburg, S. 1987. A Distributed System for Language Generation. Technical Report CMU-CMT-86-102. Carnegie-Mellon University. May.

Nirenburg, S., in preparation. DIANA: A Distributed Semantic Analyzer.

Nirenburg, S., V. Raskin and A. Tucker, 1986. On Knowledge-Based Machine Translation. Proceedings of COLING-86. Bonn, pp. 627-632.

Nirenburg, S. and V. Raskin, 1987a. The Subworld Concept Lexicon and the Lexicon Management System. *Computational Linguistics*, vol. 13, 1987, issue 3-4. (with V. Raskin).

Nirenburg, S. and V. Raskin. 1987b. The Analysis Lexicon and the Lexicon Management System. *Computers and Translation*, 2, pp. 177-188.

Passonneau, R. 1987. Situations and Intervals. ACL-87.

Pustejovsky, J. 1987 An Event Stucture for Lexical Semantics. Submitted to *Computational Linguistics*.

Pustejovsky, J. and I. Herman, 1988, Subevent Reasoning and Temporal Logic. Submission to *AAAI-88*.

Raskin, V. 1987. Linguistics and Natural Language Processing. In: S. Nirenburg (ed.). **Machine Translation: Theoretical and Methodological Issues**. Cambridge University Press, pp. 42-58.