A MODEL OF NATURAL LANGUAGE PROCESSING OF TIME-RELATED EXPRESSIONS

Yutaka Kusanagi

UNIVERSITY OF TSUKUBA
SAKURA-MURA, IBARAKI, JAPAN

This paper proposes a model of automatic processing of time-related expressions by introducing the notion of focus in the cognition level. The linguistic categories of time are determined on the basis of the relationships among the time of an extra-linguistic situation, the focused time and the time of utterance.

INTRODUCTION

It is generally understood that there is no one-to-one correspondence between the extra-linguistic (real world) time of situations and the grammatical expressions of time in sentences of natural languages. If there is no correspondence of the extra-linguistic situation to the linguistic expressions, it is impossible to simulate this process by a computer model.

This paper proposes a model of automatic processing of time-related expressions by introducing the notion of focus in the cognition level between the extra-linguistic time of situations and the linguistic expressions of time. As well, it points the way toward the computer-realization of the model. The computer programming language used for this model is SNOBOL.

Failure of correspondence between an extra-linguistic situation and linguistic expressions of the time concept can easily be demonstrated in most languages, if not all. For example, suppose the extra-linguistic situation is the speaker's writing during a certain period in the past (say, between two and three o'clock yesterday). In reference to it, we can use any of the following sentences.

- (1) I wrote a letter.
- (2) I was writing a letter.
- (3) I have written a letter.
- (4) I had written a letter (when John came).

The differences among sentences (1) through (4) are attributed to the speaker's different point of view. Comrie¹ mentions that 'aspects are dif-

ferent ways of viewing the internal temporal constituency of situations.' The speaker's way of viewing, however, affects not only his choice of aspects but also his choice of tenses. Note the following examples from Japanese.

- (5) ah sokoni aru.
'There it is!'
- (6) ah sokoni atta.
'There it was (all the time)!'

There is no difference in the denotative meaning of sentences (5) and (6). Both sentences can be used when someone has found something he has been looking for. Whether the speaker uses sentence (5) or sentence (6) depends, again, on his view point.

FOCUS

The difference in the speaker's view-point, or that of cognition, cannot be ignored, as seen in the above examples, when we consider the time-related expressions of natural language. Physical events in the extra-linguistic world present themselves as sense data, and are perceived not passively, but actively by an intervention of mind, and are subsequently stored in the memory as experiences. The experiences represent the material that can be formulated and articulated as the linguistic forms.

We believe that unless we incorporate cognition into the process of generating linguistic expressions we cannot explain the relationship between the extra-linguistic situations and the linguistic expressions. And we consider the notion of focus the most important cognitive element in connection with time-related expressions. 'Focus' is defined as the period along the extra-linguistic time line upon which the speaker directs his attention.

MODEL

By incorporating the speaker's focus into a model we can generate the time-related expressions rather

automatically, that is, by one-to-one correspondence, and therefore it is possible to implement this model as a computer model to generate linguistic expressions.

For the model we consider three levels; that is,

1. the extra-linguistic (real world) level,
2. the cognition level, and
3. the linguistic level.

Information on extra-linguistic situations observed by the speaker is stored in the extra-linguistic level in the form of the situation and its existing period of time. The speaker's focus in the cognition level is matched against the situation(s) in the extra-linguistic level and the way they match or do not match determines the selection of linguistic items such as verbs, adjectives, negation, aspect and tense.

The linguistic level forms the sentences by syntactic, morphological and phonological rules.

EXTRA-LINGUISTIC LEVEL

The computer model under consideration simulates this process by our creating three levels here also. The extra-linguistic level stores information on situations in the following form (Nos. 2-9 of the appended program).

	SITUATION	START-DS	END-DS	END-SS
SPEAK(I)	10030930	10031000		
READ(I,A-BOOK)	10021400	10021630		
COME(JOHN,HERE)	10021130	10021200	10021400	
GO-ON(A-LIGHT)	10021900	10021900	10022359	
CALM(IT,HERE)	10012000	10012000	10030500	

The information includes situation in the form of verbal expressions and their arguments, the starting time of dynamic situations, the ending time of dynamic situations, and the ending time of static situations. In the number for the time, the first two digits stand for the month, the next two for the day, the next two for the hour, and the last two for the minute. This number system may be expanded to include the year, the second, etc.

Extra-linguistic situations are divided into two categories; dynamic and static. For the distinction, let us quote Lyons⁹(p. 483) .

A static situation (or state-of-affairs, or state) is one that is conceived of as existing, rather than happening, and as being homogeneous, continuous and unchanging throughout its duration. A dynamic situation, on the other hand, is something that happens (or occurs, or takes place): it may be momentary or enduring; it is not necessarily either homogeneous or continuous, but may have any of several temporal contours;...

Dynamic situations are in turn divided into two types. Those such as 'reading' are situations occurring for a certain period of time. Therefore situations of this type have a starting time and an ending time. The second type of dynamic situation, like '(a light) going on', is momentary, and therefore it is understood that the starting time and the ending time are identical.

Static situations, on the other hand, continue for a certain period of time. But they may be results of dynamic situations and they may also change into other situations again by dynamic situations. The starting time of a static situation is usually the ending time of a dynamic situation and its ending time, if any, is the starting time of another dynamic situation.

It is worth mentioning here that what are stored in this level are not linguistic expressions but situations. Therefore, the time period of an existing situation is not influenced by the type of particular linguistic expressions of a particular natural language. Taking an English expression for example, '(a light) goes on' is considered to be momentary event. Therefore, we cannot say '*a light is going on.' However, 'lights are going on' is possible, in which case we can identify a different starting time and ending time. Further, if we could observe the period of duration of a light going on by a high-speed film for instance, the different starting and ending times could be recorded in the data by expanding time digits to second or millisecond.

However, there are some phenomena peculiar to individual languages. For instance, the situation of 'coming' is

considered to have a time duration in English. Therefore, 'John is coming' is a paraphrase of 'John is on the way.' But in Japanese it is considered to be momentary, that is, 'coming' is a sort of switch from 'is not here' to 'is here'. Therefore there is not Japanese expression equivalent to English 'John is coming.' Theoretically, this kind of difference among individual languages should be taken care of by elaborating the structure of situations in the data of this level, but in the current model it is taken care of by the linguistic level.

COGNITION LEVEL

Now in the cognition level, the focused situation and the period of focus as well as the time of utterance are inserted into the program (Program Nos. 10-27). The focused situation is in the form of verbal expression and its arguments. The period of focus is in the form of either a time adverbial such as 'now', 'this morning', 'yesterday' and 'last month', or numbers for chronological time points similar to those of extra-linguistic situation time. When the input is a time adverbial, the program calculates it against the time of utterance and translates it into chronological numbers (Program Nos. 28-40).

So far, data stored in the program are:

1. Extra-linguistic situations, each comprising:
 - a. the starting time-point of a dynamic situation(A),
 - b. the ending time-point of a dynamic situation(B),
 - c. the ending time-point of a static situation(C);
2. Time of utterance(O);
3. Focused situation, comprising:
 - a. the starting time-point of focus(S), and
 - b. the ending time-point of focus(E).

Here it is checked whether a particular extra-linguistic situation exists within the focused period or moment and, if so, how it is focused. On the basis of the relationship among the three kinds of time-points of an extra-linguistic situation (A,B and C), and the two kinds of focus time-points of the same focused situation (S and E), the categories of aspects are determined (Program Nos. 41-62).

The categories of aspects are determined in the following way:

- I. If the focus is on the Static Situation(SS);
 1. and if the focus is on the point,
 - a. and if the focus is completely inside the SS, then: STATIVE,
 - b. otherwise, the Dynamic Situation(DS) is checked.
 2. If it is otherwise,
 - a. and if the SS overlaps the focus completely, then: STATIVE,
 - b. otherwise, the DS is checked.
- II. If the focus is completely outside the DS, then: NEGATIVE.
- III. If the focus is on a point,
 1. and if the DS is on a point, then: PERFECTIVE.
 2. If it is otherwise, and
 - a. if the focus is on the beginning of the DS, then: PERFECTIVE (INGRESSIVE),
 - b. if the focus is on the ending of the DS, then: PERFECTIVE (COMPLETED),
 - c. if neither of the above, then: IMPERFECTIVE.
- IV. If the focus is on a period,
 1. and if the DS is on a point, then: PERFECTIVE.
 2. If it is otherwise, and
 - a. the DS overlaps the focus completely, then: IMPERFECTIVE,
 - b. if the DS overlaps the focus partially, and
 - i. the beginning of the DS is on or before the ending of the focus, then: PERFECTIVE (INGRESSIVE),
 - ii. the ending of the DS is on or before the ending of the focus, then: PERFECTIVE (COMPLETED),
 - iii. it is neither of the above, then: PERFECTIVE.

This aspect assignment is illustrated in the diagrams in the Appendix I.

The categories of tenses are in turn determined on the basis of the relationship between the focused time and the time of utterance (Program Nos. 63-69). If the focus is on the time of utterance, the PRESENT is assigned. If the focus is before the time of utterance, the PAST is assigned; if the reverse, the FUTURE. Japanese and Russian, for example, work in this way. In English one more

category of tense is needed. That is, if the focus is on a period and the end of the focus coincides with the time of utterance, the PAST-PRESENT is assigned. This category will later cover expressions like 'I have been studying linguistics.'

FINAL REMARKS

In the linguistic level time-related morphemes are assigned on the basis of the categories of time-related expressions sent from the cognition level. Here the peculiarities of individual natural languages are involved. Therefore a unique program is required for each natural language. After the morpheme assignments, syntactic and morphological rules together with phonological rules determine the forms in the sentences.

We shall omit detailed discussion on this level, since this process involving word categories closely related to the distinction between dynamic and static situations is complicated. For simplicity of explanation, we have omitted the discussion and the programming of expressions of habit, experience and atemporal expressions. We have also omitted the discussion of relative tenses.

In summary, we have discussed an automatic process of time-expression generation in natural language. We believe that this is only possible by the introduction of the speaker's time focus and an examination of the relationship between the extra-linguistic situational time and focused time. By further examining the manner of morpheme assignments in as many natural languages as possible, we believe, we can expand the number of universal rules common to all languages and minimize the number of rules peculiar to individual languages.

(This research is partially supported by Japanese Ministry of Education grants Nos. 361200 and 561151.)

REFERENCES

1. Comrie, B. (1976). Aspect. Cambridge: Cambridge University Press.
2. Forsyth, J. (1970). A Grammar of Aspect. Cambridge: Cambridge University Press.

3. Kusanagi, Y. (1972). "Time Focus within the Japanese Tense System", Papaers in Japanese Linguistic 1-1, 52-68.
4. Kusanagi, Y. (1975). "Gengo Katsudo ni Okeru Ninchi Sayo--imiron ni Okeru Ichi Kasetu (Cognition in Linguistic Behavior: a hypothesis in semantics)", Gengo no Kagaku (Sciences of Languages) 6, 85-112.
5. Lyons, J. (1977). Semantic II. Cambridge: Cambridge University Press.

APPENDIX I

Diagrams of Aspect Assignment

-
- A -- starting time of a dynamic situation
 B -- ending time of a dynamic situation
 C -- ending time of a static situation
 S -- starting time of focus
 E -- ending time of focus
-

	PERFECTIVE (COMPLETED)		PERFECTIVE (COMPLETED)
	PERFECTIVE (COMPLETED)		NEGATIVE
	PERFECTIVE		STATIVE
	IMPERFECTIVE		STATIVE
	IMPERFECTIVE		NEGATIVE
	PERFECTIVE		STATIVE
	IMPERFECTIVE		STATIVE
	NEGATIVE		STATIVE
	PERFECTIVE (INGRESSIVE)		NEGATIVE
	NEGATIVE		PERFECTIVE (COMPLETED)
	PERFECTIVE (COMPLETED)		PERFECTIVE (COMPLETED)
	PERFECTIVE		PERFECTIVE (COMPLETED)
	NEGATIVE		PERFECTIVE (COMPLETED)
	NEGATIVE		PERFECTIVE
	IMPERFECTIVE		PERFECTIVE
	STATIVE		PERFECTIVE
	STATIVE		PERFECTIVE
	NEGATIVE		PERFECTIVE (COMPLETED)
	STATIVE		PERFECTIVE (COMPLETED)
	NEGATIVE		PERFECTIVE
	STATIVE		PERFECTIVE
	STATIVE		*****

APPENDIX II

```

@ASNOB,E
ASCII SNOROL 74R1.1 06/06/80 09:33:56
EXPERIMENTAL VERSION 2 200531
1  DATE = 'TODAY' ! 'YESTERDAY' ! 'TOMORROW'
2  DCHK1 CHECK = TRIM(INPUT)
3  CHECK = 'E'
4  ELDATA DATA = TRIM(INPUT)
5  DATA = '*'
6  N = N + 1
7  DATA ARB = $( 'ELST' N ) ' ' ARB = $( 'A' N ) ' ' REM
8  RE ARB = $( 'B' N ) ' ' REM = $( 'C' N )
9  RE ARB = $( 'B' N ) ' ' REM = $( 'C' N )
10 DCHK2 CHECK = 'S'
11 SDATA DATA = TRIM(INPUT)
12 DATA = '*'
13 SPTIME = DATA
14 O = SPTIME
15 SPTIME LEN(2) . MO LEN(2) . DAY LEN(2) . HR LEN(2) . MNT
16 FDATA DATA = TRIM(INPUT)
17 DATA = '*'
18 DATA ARB = FST ' ' REM = TADV
19 K = 0
20 KADD K = K + 1
21 FST = ' '
22 $( 'ELST' K ) FST
23 EQ(N,K)
24 ABC A = $( 'A' K )
25 ST = $( 'ELST' K )
26 B = $( 'B' K )
27 C = $( 'C' K )
28 DATE TADV DATE
29 TADV = 'TODAY'
30 ADAY = DAY + 00
31 D1 TADV = 'YESTERDAY'
32 ADAY = DAY - 01
33 D2 TADV = 'TOMORROW'
34 ADAY = DAY + 01
35 MO TADV = 'MONTH'
36 SE TADV LEN(8) . S ' ' LEN(8) . E
37 ADJUST LT(ADAY,10)
38 ADAY = '0' ADAY
39 AGO S = MO ADAY '0000'
40 E = MO ADAY '2359'
41 S0 C LEN(1)
42 S1 GE(S,C)
43 S2 LT(S,B)
44 S3 EQ(S,B)
45 S4 EQ(E,B)
46 S5 LE(E,C)
47 C1 LT(E,A)
48 C2 GT(S,B)
49 C3 EQ(E,A)
50 C4 EQ(A,B)
51 C5 EQ(B,S)
52 C6 EQ(A,S)
53 C7 LE(A,B)
54 C8 GT(E,B)
55 C9 GE(E,B)
56 NEG ASP = 'NEGATIVE'
57 IMP ASP = 'IMPERFECTIVE'
58 PEF1 ASP = 'PERFECTIVE (COMPLETED)'
59 PEF ASP = 'PERFECTIVE'
60 PEF2 ASP = 'PERFECTIVE (INGRESSIVE)'
61 ST1 ASP = 'STATIVE'
62 ST2 ASP = 'STATIVE (COMPLETED)'
63 *** JAPANESE TENSE ASSIGNMENT ***
64 TN EQ(S,E)
65 T1 EQ(S,O)
66 T2 LT(E,O)
67 T3 GT(S,O)
68 PRE TEN = 'PRESENT'
69 PST TEN = 'PAST'
70 FUT TEN = 'FUTURE'
71 TA
72 FST = FST ' '
73 OUTPUT = 'FOCUS ON ' FST ' FROM ' S ' TO ' E
74 OUTPUT = 'EX-L SITUATION '
75 OUTPUT = ' ' ST
76 OUTPUT = ' ' A ' A
77 OUTPUT = ' ' B ' B
78 OUTPUT = ' ' C ' C
79 OUTPUT = 'EXPRESSION (ASPECT) ' ASP
80 + (TENSE) TEN
81 OUTPUT = '*****' : (FDATA)
82 END

```

NO ERRORS WERE DETECTED IN SOURCE PROGRAM.

 FOCUS ON SPEAK(I) FROM 10030940 TO 10030940
 EX-L SITUATION
 SPEAK(I)
 A 10030930
 B 10031000
 C
 EXPRESSION (ASPECT) IMPERFECTIVE (TENSE) PRESENT

 FOCUS ON READ(I) FROM 10020000 TO 10022359
 EX-L SITUATION
 READ(I,A-BOOK)
 A 10021400
 B 10021630
 C
 EXPRESSION (ASPECT) PERFECTIVE (TENSE) PAST

 FOCUS ON COME(JOHN) FROM 10020000 TO 10022359
 EX-L SITUATION
 COME(JOHN,HERE)
 A 10021130
 B 10021200
 C 10021400
 EXPRESSION (ASPECT) PERFECTIVE (TENSE) PAST

 FOCUS ON CALM) FROM 10020000 TO 10022359
 EX-L SITUATION
 CALM(IT,HERE)
 A 10012000
 B 10012000
 C 10030500
 EXPRESSION (ASPECT) STATIVE (TENSE) PAST

 FOCUS ON GO-ON(A-LIGHT) FROM 10022000 TO 10022359
 EX-L SITUATION
 GO-ON(A-LIGHT)
 A 10021900
 B 10021900
 C 10022359
 EXPRESSION (ASPECT) STATIVE (TENSE) PAST