

# A Novel Analysis of Temporal Frame-Adverbials

Magnus MERKEL

Department of Computer and Information Science

Linköping University

S-581 83 LINKÖPING

SWEDEN

Telephone: +46 13282423

Internet: MME@IDA.LIU.SE

## Abstract

In this paper interpretation principles for simple and complex frame-adverbial expressions are presented. Central to these principles is a distinction between *phases* and *periods* together with the *temporal hierarchy*, where multiple scales of time and relations can be expressed. A system, CLOCKWISE, has been implemented which interprets Swedish temporal expressions according to the principles outlined in the paper.

## Introduction

Temporal information is expressed and conveyed in a number of ways in natural language including tense, aspect and lexical items that carry temporal information, eg. temporal adverbs. Most researchers in this field approach temporal entities in language from the perspective of tense and aspect. But there is very little in the literature on other expressions that hold temporal information, such as temporal adverbs, certain prepositional phrases and noun phrases. In most papers the meaning of a temporal adverbial such as 'next year' is merely explained as the predicate 'next year' which specifies a point or interval of time from a reference time. In objective time-modelling systems such as /Kahn & Gorry 1977/ and /Bruce 1973/ temporal expressions were never analyzed in their linguistic form; instead they had to be typed in as stereotyped lists. The internal structure of temporal expressions must be investigated in order to construct grammars that can capture general features and be of use in computational applications. In the paper I will focus on temporal frame-adverbial phrases, that is, expressions that refer to a temporal period in which events are located (Cf /Bennet & Partee 1978/, /Hinrichs 1986/). I will not discuss the complex question of how tense, aspect and temporal adverbials interact (for an outline of the problems see /Ejerhed 1987/).

## Temporal frame-adverbial phrases

A large group of temporal expressions can be classified as frame-adverbial phrases. /Smith 1981/ categorizes temporal frame-adverbial phrases in the following way<sup>1</sup>:

Deictic	Clock-Calendar <sup>2</sup>	Dependent
- last week, yesterday	- at midnight	- previously, before
= now, this moment		= the same time
+ next week, tomorrow	+ at midnight	+ later, afterwards

<sup>1</sup> My use of + and - is the same as Smith's forward and backward arrows.

<sup>2</sup> In an earlier paper (Smith 1980) Smith calls this class *flexible anchoring adverbials*.

This classification is based on the concept of anchoring, ie. a deictic is generally anchored to the time of speech (ST), a dependent is anchored to another given time in the context, and a clock calendar adverbial can anchor to either ST or to some other context-dependent time. Smith attributes a relational value to these expressions, -, = and +. These symbols stand in turn for the relational value *anteriority* (past), *simultaneity* (present) and *posteriority* (future).

Frame-adverbials can syntactically occur as adverbs, noun phrases and prepositional phrases. They can be complex expressions such as

(1) *On Monday next week at 6 pm.*

where temporal information is specified on several levels and we have a combination of deictic and clock calendar expressions. Complex expressions have a relatively loose syntax in that the time denoted in (1) can be expressed as (2) and (3)

(2) *Next week on Monday at 6 pm.*

(3) *At 6 pm next week on Monday.*

The issues that need considering are the following: How is the semantic well-formedness of frame-adverbial phrases determined and, if possible, what is needed to establish their temporal reference in terms of locations on the time axis.

## Periods and phases

To account for the semantic well-formedness of frame-adverbial phrases the 'time' denoted by the phrase can be thought of as an ordered set of specifications for conventional temporal concepts, such as YEAR, MONTH, DAY, HOUR, MINUTE, SECOND. The ordering of these temporal concepts may seem to fit well into a 'Chinese box metaphor' where each concept is included in a concept on a higher level and in turn includes a concept on a lower level. A semantically well-formed temporal frame-adverbial would then be described as a phrase which forms a description of an unbroken chain of temporal concepts. This is put forward by /Hinrichs 1986/ in his *Scoreboard of reference points* and *cooccurrence restriction* in which the construction of a reference time from a frame-adverbial phrase is possible if there is a chain of specified temporal concepts in accordance with an ordered set of concepts. However, if we included WEEK in the above set of concepts we would get into trouble. Weeks are not included in months in the same way as months are included in years and a problem of where to fit in WEEK in the Chinese box order would occur. This is due to the fact that the western calendar system is based on three separate descriptive traditions: the week, the Gregorian calendar and the clock /Levy 1980/. The Gregorian calendar holds year, month and day(of month) as its basic concepts, the week is a sequence of named days, and the clock system divides the day into hours and hours into seconds, etc. The week and Gregorian calendar share the concept of day and the clock

system splits the day into time segments. The reason that we can express a time by specifying different concepts of separate systems is that the systems share one or more concepts with each other. It is also reflected in the way we refer to the day period (in Swedish and English). A *day* can be referred to in two different ways, either by specifying the day of a certain week (DW) or by picking out a day of a certain month (DATE). In English the expression *Monday next week* will describe the same day on the time axis as an expression describing the same day by specifying a year, a month and a day of month. In Swedish it is possible to specify a unique day by the year-week-day path as the Swedish calendar has numbered weeks, giving us a year divided into 53 weeks<sup>3</sup>.

In our model we want to distinguish between *periods* and *phases* and show the relations between them.

A period *P* is a segment of time with a certain length, such as YEAR, MONTH, DAY, WEEK, HOUR, MINUTE, SECOND.

A period *P<sub>i</sub>* is a subperiod of a period *P* iff *P* can be seen as consisting of a number of periods *P<sub>i</sub>* that together exhaust *P*.

A phase *p* is a specific instance of a subperiod having a certain order in the sequence of subperiods exhausting a given period *P*.

For example, a period such as YEAR can be seen as consisting of twelve subperiods (MONTH). *January* is the first phase of those twelve subperiods.

The graph below illustrates the relation between year as a period and the monthly phases.



Fig. 1 The YEAR period

This graph shows the year period from one point of view. We could split the year into other phases, for instance, into periods of seasons, where summer, autumn, winter and spring would be the corresponding phases. Note that a sequence of phases is cyclically recurring.

Apart from differentiating periods and phases we must be able to show the relations between them. The following table shows some periods with corresponding phases:

Periods	Subperiod of	Phases
YEAR:	CALENDAR	{...0...}
MONTH:	YEAR	{January,...December}
WEEK:	YEAR	{1, 2, ...53}
DAY:	MONTH	{1, 2,...31}
	(DATE)	
HOUR:	WEEK	{Monday,...Sunday}
	(DW)	
MINUTE:	DAY	{0, 1,...23}
SECOND:	HOUR	{0, 1, ...59}
	MINUTE	{0,1, ...59}

<sup>3</sup> One could say that the system of numbered weeks is similar to the use of named months or named days. The analogy does not hold all the way of course as the week system is historically independent from the Gregorian calendar, but we have 'tried' to force the week concept onto the concept of year by naming the weeks with numbers. In Sweden a numbered week always starts on a Monday. The first week that holds at least four days of a new year forms week 1. The first of January 1987 was a Thursday and part of week 1, whereas the first of January 1988 is a Friday which results in the first three days of 1988 being part of week 53 1987! This is standardized by Swedish authorities.

For each period there are one or two possibilities to view it as a subperiod of a higher-level period. The difference between the two different ways of referring to a DAY period is shown by the fact that there are two *subperiod-of* relations, one to MONTH, and one to WEEK.

The *subperiod-of* relations can be displayed in a *temporal hierarchy* of the following kind.

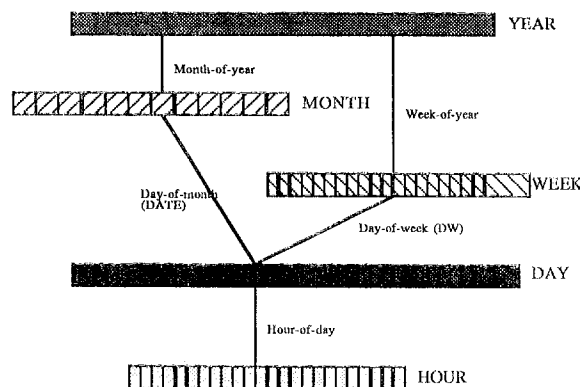


Fig. 2 The temporal hierarchy

The lines connecting temporal periods to each other show subperiod-of relations. The hierarchy shows the two different paths from YEAR to DAY.

The periods and phases described above share one property and that is that they all have fixed lengths and as a consequence *clear boundaries*. We can easily locate the point on the time axis where *January* becomes *February*, but we cannot with accuracy specify the location on the time axis where spring is succeeded by summer. So we must distinguish between *bounded* and *fuzzy* periods. There are other phasal divisions one can make. For example, we could split the week into weekdays and non-weekdays (week-ends?), and the year into feasts, such as Christmas and Easter, and non-feasts, etc. It is just a matter of fact that we divide time into different temporal scales, and some contain periods of fixed length and some do not. This is why it is problematic to try and map a phase of no fixed length, such as *summer*, onto another temporal scale where we are dealing with fixed boundaries such as the one depicted in the temporal hierarchy above.

Durative expressions are mostly expressed in terms of quantified periods: *one week, two months, three years*, etc. It is more seldom that phasal expressions are used in a durative sense. Quantified phasal expressions such as *two mornings, six Thursdays, three winters*, etc are primarily used to express habituality and iteration. It is however possible to infer duration from these as we know the length of a phase or can infer it from its corresponding period.

### Phasal and deictic expressions

In accordance with the previous section we can now distinguish between *phasal* and *deictic expressions*.

In our terminology a phasal expression is a temporal expression that primarily describes a phase by using a corresponding word for the phase in question. A phasal expression may be a single word (*1988, January, midnight*) or be more complex, such as a prepositional phrase (*in 1988, in January, at midnight*).

The difference between Smith's clock calendar adverbials /Smith 1981/ and phasal expressions is that phasal expressions are just "providers of phases" and they are not primarily anchored to any other time as clock calendar adverbials are. Phasal expressions may describe absolute times if they refer to a unique phase on the time axis. So if the phasal expression 1987 denotes a *unique* phase of the calendar then it is possible to map it into an absolute time. Following /Hinrichs 1986/ we call these phasal expressions *complete or independent dates*. Also included in this class we find complex phasal expressions which, if mapped on to the temporal hierarchy, would form a *chain of temporal information*. By chain of temporal information we mean a chain of temporal phases according to the temporal hierarchy.

Deictic expressions function basically in the following way:

Starting from a reference time (eg. ST) a deictic will establish a new reference time with a different phasal value for the relevant period by moving a number of steps in either a forward (+) or a backward (-) direction depending on the deictic, and keeping phasal values for higher-level periods consistent with this change.

Certain deictics do not change the reference time. For example, *i år (this year)*, *den här veckan (this week)*, *i dag (today)*, etc, do not move the reference time in a forward or backward direction. The new RT will be a partial copy of the old RT; *today* will describe a time equal to ST and specified to the DAY period level.

The deictic expression *nu (now)* is ambiguous in respect of period level. In one context *nu* would map onto the year level and in another context it would refer to the hour level of RT.

In our system the function that maps a deictic expression is of the form  $F(x, index, step)$  where  $x$  is either a phase or a period. Applying  $F$  results in a *description* of the time in terms of a set of phases which forms a chain in the temporal hierarchy with  $x$  as its most specific phase or period. Posterior deictics will have positive numbers as *step*, anterior will have negative numbers and simultaneous will have 0. For example, a deictic such as *i morgon (tomorrow)* will be mapped to a set of corresponding phases from the YEAR period down to the DAY period, leaving more specific phases unspecified. This means that while constructing a description of a time unnecessary work is avoided as we are dealing with partial descriptions and not the temporal objects as such. Compare the following two phrases:

- (1) *tomorrow*
- (2) *tomorrow at 3 o'clock.*

In both expressions *tomorrow* will contribute with exactly the same information, ie. a *description* of a certain day of a certain month in a certain year, etc.

Deictics in Swedish can be categorically characterized in the following way:

- Lexical deictics (such as *igår (yesterday)*, *fjol (last year)* *idag (today)*).
- Pre/postpositional phrases with a temporal NP where the pre/postposition signals the type of temporal expression. For example, *om två veckor (in two weeks)*, *för fem år sedan (five years ago)*.
- Noun phrases with a temporal phase or period word as head and a modifier or ordinal that signals a deictic function, such as (*nästa vecka (next week)*, *förra året (last year)*), etc.

In many contexts prepositions are omitted in Swedish deictic expressions if the remaining NP signals a deictic function on its own.

Note also that a phasal expression such as *in October* may, apart from specifying a phase *October*, also function deictically, that is, it may provide phasal values for the periods above its explicit level. However, this is regarded as a secondary use of a phasal expression.

Determining the initial reference time (or index) for deictic expressions is another important consideration. Some deictic expressions are necessarily anchored to ST and others may be anchored to ST but they may also be anchored to another reference time depending on context. In Swedish expressions such as *igår (yesterday)*, *i morgon (tomorrow)*, *om 2 dagar (in two days)* and *förra året (last year)* are examples of expressions that must anchor to ST. Some expressions involving *nästa (next)* are examples of the second group. The expressions *nästa vecka (next week)* and *nästa torsdag (next Thursday)* can have either ST or another context-dependent time as their initial reference time. Expressions such as *nästa dag (next day)* and *två veckor senare (two weeks later)* must necessarily anchor to a reference time other than ST (see /Merkel 1987/).

### Next-expressions

Expressions such as *nästa torsdag (next Thursday)*, *nästa sommar (next summer)* and *nästa jul (next Christmas)* may be ambiguous to certain speakers. These expressions seem to be ambiguous in the sense that they either take you 1 or 2 steps from your reference time. So *nästa torsdag* may refer to the first or second *Thursday* from the initial reference time. If we changed the argument of *nästa* in the above examples to words denoting periods instead of phases the corresponding reference times would not be ambiguous. *Nästa vecka (next week)* and *nästa dag (next day)* will be interpreted as having only one temporal location and not two as the first examples. How is this possible?

If somebody says *nästa torsdag (next Thursday)* when the initial reference time is a Thursday, then there is no ambiguity; the step value will be +1 which will pick out the first phase that has the value *torsdag* after the initial RT. However, if *nästa torsdag* was uttered on a Sunday, we would have to adjust the initial RT by either starting at the Thursday preceding the initial RT or at the Thursday succeeding it, and from the 'new' initial RT move +1 to reach a temporal reference. The expression *nästa jul (next Christmas)* is analogous to *nästa torsdag*. Instead of saying that *next-expressions* have two possible step values, +1 or +2, we claim that there is only one value, namely +1, but that when the argument is a phasal expression, such as *torsdag*, the initial RT may be chosen in two ways depending on individual strategies. Some may for example choose the nearest phase as their new RT.

There is a class of expressions that behaves the same way as next-expressions, for example, expressions involving modifiers such as *förra (= last)*, *nästinästa (the next but one)*.

### Interpretation principle

By looking at the properties of the different parts of a complex temporal expression interpretation principles for frame-adverbial expressions can now be set up. Making use of the distinction between the deictic and phasal expressions together with the temporal hierarchy we can outline a Frame-adverbial interpretation principle which will explain when a frame-adverbial expression is interpretable:

#### Frame-adverbial interpretation principle:

It is possible to interpret a temporal frame-adverbial phrase autonomously iff

1. a) The phrase consists of one or several phasal expressions  
and  
b) The combination of phasal values will constitute a chain of temporal information according to the temporal hierarchy.  
and  
c) The top-most phase of the chain specifies a unique phase.

or

2. a) The phrase consists of a combination of one or several deictics and an arbitrary number of phasal expressions  
and  
b) Index is known for each deictic  
and  
c) The deictics and the phasal expressions each specify consistent phase-of-relations. The phasal values provided will constitute a chain of temporal information according to the temporal hierarchy.  
and  
d) A deictic provides temporal information for the highest level of the chain.

The above principle will permit interpretation of expressions such as

- |   |   |
|---|---|
| (1) 1987 i maj<br>den 24 maj 1987   | 1987 in May<br>May 24th, 1987   |
| (2) i morgon kl kvart över fyra<br>nästa vecka på torsdag kl<br>kvart över fyra<br>för 20 år sedan idag<br>vid den här tiden i morgon<br>för 39 år och fyra månader sedan<br>om två veckor och fyra dagar | tomorrow at a quarter past four<br>next week on Thursday at<br>a quarter past four<br>20 years ago today<br>at this time tomorrow<br>39 years and four months ago<br>in two weeks and four days |

The following expressions will be ruled out by the principle:

- |  |  |
|--|--|
| (3) kl kvart över 5 1987<br>på kvällen i januari     | at a quarter past five 1987<br>in the evening in January   |
| (4) förra året på eftermiddagen<br>nästa vecka kl 12 | last year in the afternoon<br>next week at twelve o' clock |

The fact that an expression is ruled out by the principle does not mean that interpretation is blocked completely. The expressions in (3) and (4) are only incomplete in the sense that some periods lack a phasal value and it may turn out that contextual knowledge will provide this. Interpretation is stuck if the expression yields inconsistent phasal values. For example, the expression *på måndagen den 23 november i år* (*on Monday November 23rd this year*) provides inconsistent phasal values for the DAY period, ie. November 23rd is a Wednesday in 1988.

However, there are some borderline cases that the above principle would permit such as

- |  |   |
|--|---|
| (5) för 39 år sedan igår<br>nästa vecka nästa år<br>för 2 år sedan för 2 månader sedan | 39 years ago yesterday<br>next week next year<br>two years ago two months ago |
| (6) för 39 år sedan i morgon<br>om sex år för 2 månader sedan                          | 39 years ago tomorrow<br>in six years 2 months ago                            |

In (5) two deictics with the same direction are combined, either + +, or - -. And in (6) there is a combination of deictics with different relational values, that is an anterior is joined with a posterior deictic. In principle it seems possible to combine deictics in this way, although some Swedish speakers do consider the examples in (5) and (6) as being odd.

The examples in (5) and (6) are all complex frame-adverbial expression holding more than one deictic phrase, but the important thing is that the different deictics *together* specify the same time. In other words, if we interpret a phrase such as

(7) *för två år sedan igår* (*two years ago yesterday*)

where the two involved deictics both have ST as obligatory index the first deictic will pick out a description of a reference time two years before ST and the second will modify that description by adding phases for DAY and MONTH based on the partial description of a phase one day before ST. This view makes it possible to claim that certain deictics such as *igår* (*yesterday*), *i morgon* (*tomorrow*), etc, *always* have ST as their index. /Smith 1978/ has proposed that interpretation of expressions such as (7) is done in two steps, that is, that the first deictic forms the RT for the second. Smith's strategy would mean that *igår* (*yesterday*) in (7) cannot have ST as its obligatory index; with our strategy this is indeed possible.

#### Implementations

The principles outlined for the interpretation of temporal frame-adverbials have been implemented in a system, CLOCKWISE, which interprets frame-adverbial expressions into a temporal description in terms of phasal values. The first version of CLOCKWISE consists of a parser, based on finite-state machinery, and a 'temporal expert' that will make use of its knowledge about temporal phases and periods and infer temporal information that is missing explicitly in the expressions. A temporal representation is constructed in a *notepad* during the parsing process. The notepad contains information about the periods and phases denoted by the expression and, if the expression is deictic, also index, step value and direction. The notepad is structured according to the temporal hierarchy and will therefore support the interpreter according to the Interpretation principle. The first version can, however, only deal with one type of index, namely speech time. CLOCKWISE 1 has been used as a module in a natural language and graphical interface to a booking system /Jönsson 1987/.

The second version of CLOCKWISE is under construction. The temporal representation is built by an LFG-type grammar together with a frame-based knowledge representation where temporal periods and phases are treated as semantic objects (cf /Ahrenberg 1988/). The grammar is written in a way to syntactically filter out the temporal expressions, ie. temporal prepositional phrases are functionally distinct from other prepositional phrases in the functional structures (f-structures). The result of a successful interpretation of a temporal frame-adverbial is a directed acyclic graph (DAG) where a set of phasal values forms a description of the time denoted by the expression. For example, a frame-adverbial such as *nästa vecka på fredag kl 12.30* (*next week on Friday at 12.30*) has a functional structure of the following kind:

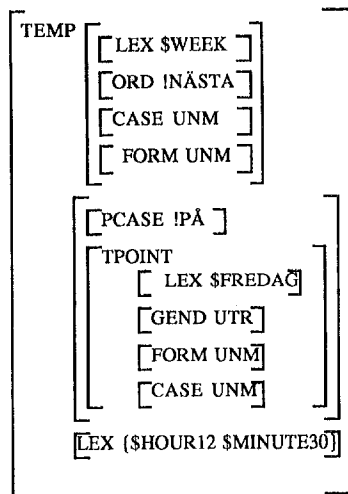


Fig. 3 Functional structure

The above DAG is mapped to an intermediate DAG by the Frame system. The intermediate DAG has separated deictic phrases from phasal ones, and also provided a frame for the deictic mapping function *F*.

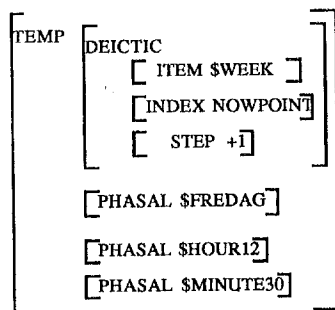


Fig. 4 Intermediate structure

By using a more general and powerful framework we are able to incorporate other types of temporal expressions and develop the above principles in relation to tense and aspect. Durative expressions have recently been incorporated in the system. The development of CLOCKWISE is part of the grammar development project in the LINLIN project at Linköping University (Ahrenberg 1987).

CLOCKWISE is at the moment unable to handle expressions such as

(1) *The first week of January*

CLOCKWISE will get stuck on this expression due to the fact that weeks cannot be mapped onto months, whereas *the first day of January* would result in a description. One solution would be to force a mapping of weeks onto months (in the same way as weeks are mapped onto years in the Swedish calendar) giving us five subperiods of MONTH each of week length, with corresponding phases where the first and the fifth week phase sometimes would be partial.

## Acknowledgements

This research has been supported by the National Swedish Board for Technical Development.

I would like to thank Lars Ahrenberg, Nils Dahlbäck, Arne Jönsson, Mats Wirén och Ivan Rankin for helpful comments on drafts of this paper. I am also indebted to Jim Goodwin who has provided me with the knowledge representation tool used in CLOCKWISE 2.

## References

Ahrenberg, L, 1987, Parsing into Discourse Object Descriptions, in *Proceedings of the Third Conference of the European Chapter of the Association for Computational Linguistics*, Copenhagen.

Ahrenberg, L, 1988, Functional Constraints in Knowledge-Based Natural Language Understanding, in *Proceedings of the 12th International Conference on Computational Linguistics*, Budapest.

Bennett, M & Partee B, 1978, *Toward the Logic of Tense and Aspect*, Indiana University Linguistic Club, Bloomington.

Bruce, B, *The processing of time phrases in CHRONOS*, Report CBM-TM-29, Department of Computer Science, Rutgers University, New Brunswick.

Ejerhed, E, 1987, *Event structures in text*. Paper presented at the Workshop on Text and courses of events, held at the Department of Computer & Information Science, Linköping University, March 12-13, 1987.

Hinrichs, E, 1986, Temporal anaphora in discourses of English, in *Linguistics and Philosophy* 9, pp 63-82.

Jönsson, A, 1987, *Naturligt språk för användardialog och databasförfrågningar*. Report LiTH-IDA-R-87-25, Department of Computer & Information Science, Linköping University.

Kahn, K & Gorry G A, Mechanizing Temporal Knowledge, in *Artificial Intelligence* 9 (1977), pp 87-108.

Levy, D M, 1980, *The architecture of text*, Ph.D. dissertation, Stanford University.

Merkel, M, 1987, The Interpretation of Swedish Temporal Frame-Adverbial Phrases, in *Papers from the Tenth Scandinavian Conference of Linguistics*, Bergen.

Smith, C. S. 1978, The syntax and interpretation of temporal expression in English. *Linguistics & Philosophy*, 2, pp 43-100.

Smith, C. S., 1980 Temporal Structures in Discourse, in *Time, Tense and Quantifiers*, (ed C Rohrer), Niemeyer, Tübingen, pp. 355-374.

Smith, C. S, 1981, Semantic and Syntactic Constraints on Temporal Interpretation, in Tedeschi, P J & Zaenen, A (eds), *Syntax and Semantics, Volume 14*, pp 213-238.