

## ÜBER ZEITREFERENZ UND TEMPUS

Dieter Wunderlich

1 Berlin 41 - Steglitz  
Brentanostrasse 67

Abstract : On time reference and tense.

The problems of time reference and tense in natural languages are discussed in their syntactic and semantic aspects. The syntactic description is based on the principles of the generative transformational grammar. Since tense is a deictic category, the semantics language must be a pragmatically extended system. Furthermore, it seems that the use of tense and time adverbials in languages of highly civilized cultures can be adequately described in terms of an extensional language. (In all these cultures, conflicts concerning time reference are solved by consulting an extensional clock and calendar system.) The semantic description is based on work done by the logicians Montague and Kamp from UCLA. In addition to the theoretical treatment of problems of time reference, the possible applications of the proposed system will be sketched.

### The syntactic system.

Tense is considered not as a constituent of a base-phrase-marker, but as a syntactic feature as condition for inflexional rules. These features are attached to sentence constituents. They are to be understood as binary relations [e rel s] between reference indices which can receive a definite semantic interpretation; with rel = {before(vor), after(nach) overlapping(übl)} ; one index is called the index of event time e, the other the index of utterance time s.

Unlike the tense, time adverbials are introduced as constituents of a base-phrase-marker. They are subcategorized into features similar to the tense features. It is discriminated between adverbials which are related to context ( Advb<sub>k</sub> → [e rel e] ) and those which are related to the time of utterance ( Advb<sub>s</sub> → [e rel s] ) Since the latter imply corresponding tenses they are treated as alternatives to the tenses. For the rules cf. p. 5/6.

### The semantic system.

The semantic system is a system of predicate logic with the following descriptive extension :

- a set of variables over time intervals t, a primitive ordering relation "before" , together with some axioms defining the properties of "before". Each variable might be understood as a possible argument of sentences in the ordinary predicate logic.
- Several models for the time system are specified : the time intervals are represented as elements of the power set of the real numbers, or the time intervals are represented as real numbers with neighbourhoods (stating a topology over the set of real numbers). Sometimes an approximation of the minimal time intervals to time points is used, too. In this case time intervals simply are represented as intervals of time points.

The translation from syntactic to semantic system is obtained by means of the reference indexes.

- Replace each symbol  $e$  or  $s$  by an element of the set of variables over time intervals.
- Replace each syntactic feature of the form  $[e \text{ rel } s]$  or  $[e \text{ rel } e]$  by a set of pairs  $\{\langle T_e, t_s \rangle\}$  resp.  $\{\langle T_e, t_e \rangle\}$ , in which the first element is a variable over a set of possible event times and the second element a variable  $t$  resp.  $t$  ranging over a given set of utterance times resp. event times. The symbol 'rel' in the feature causes some specified conditions which hold between  $T_e$  and  $t_s$  resp.  $T_e$  and  $t_e$ .

#### The time structure of texts.

Some preliminary considerations about possible basic time structures of texts are given. The following notion for the time structure TS of a text containing  $m$  sentences is proposed.

$TS_m = \langle UTV, EOR, Id, R \rangle$ , with UTV = immediate text predecessor, EOR = events ordering relation, Id = Identity, R = matching from the totally ordered set of sentences into the ordered set of events.

#### 1. Ziel der Untersuchung.

Mit den folgenden Ausführungen soll in erster Linie ein Beitrag zur Grammatiktheorie natürlicher Sprachen geleistet werden. Erst in zweiter Linie und vorläufig noch sehr fragmentarisch werden Überlegungen angestellt, wie auf der Grundlage der dargestellten Theorie Algorithmen entwickelt werden können zur Analyse von Texten und zur Rekonstruktion der in ihnen beschriebenen Ereigniszusammenhänge.

Es wird versucht, ein formales System zur Repräsentation von zeitreferentiellen Beziehungen zu entwickeln. Diese werden sowohl unter syntaktischem wie auch unter semantischem Gesichtspunkt untersucht. Zeitbezug ist in den natürlichen Sprachen durch eine Anzahl verschiedenartiger Ausdrücke möglich: durch Zeitadverbien, Zeitkonjunktionen und präpositionale Zeitbestimmungen (sie werden zusammengefaßt unter dem Namen Zeitadverbial), ebenfalls durch verbale und nominale Ausdrücke (die hier unberücksichtigt bleiben, soweit nicht die nominalen Ausdrücke als Reduktionsformen aus Sätzen aufzufassen sind). Außerdem ist impliziter Zeitbezug z.B. durch Kausalbestimmungen möglich. Ferner werden die Tempusmorpheme zum Ausdruck von Zeitbeziehungen verwendet (allerdings können sie außerdem auch modale und aspektuelle Beziehungen ausdrücken). Zeitadverbiale und Tempusmorpheme werden gemeinsam unter dem Namen Zeitausdrücke angeführt.

Für die syntaktische Repräsentation der Zeitausdrücke wird ein System von Regeln zugrundegelegt, das weitgehend den Postulaten der generativ-transformationellen Grammatiktheorie entspricht. Bei der Formulierung des semantischen Systems wird auf Vorschläge der kalifornischen Logiker Montague und Kamp zurückgegriffen 2).

Es werden durchweg Verhältnisse der deutschen Sprache zugrundegelegt; doch ist das Verfahren ohne Schwierigkeit auf verwandte Sprachen zu übertragen. Es kann sogar vermutet werden, daß das Beschreibungssystem auf alle die Sprachen angewendet werden kann, in denen ein expliziter Bezug auf externe Zeitskalen (d.h. auf Uhren- und Kalendersysteme) möglich ist, damit praktisch auf alle Zivilisationssprachen. Da in diesen Sprachen Konflikts- oder Zweifelssituationen hinsichtlich einer Zeitreferenz gemeinhin dadurch entschieden werden, daß auf das konventionelle Uhren- und Kalendersystem verwiesen wird (das von extensionalem Charakter ist), scheint es gerechtfertigt zu sein, auch für das semantische Beschreibungssystem extensionale Modelle anzugeben.

Anm.1) Die nachfolgenden Abschnitte 2 und 3 enthalten einen Auszug aus meiner Dissertation "Tempus und Zeitreferenz im Deutschen", TU Berlin 1969.

2) Richard Montague, "Pragmatics and Intensional Logic". Southern California Logic Colloquium, 6. Januar 1967 (Mimeogr.).

J.A.W.Kamp, "On tense logic". Logic Colloquium UCLA, 20. Febr. 1967 (mimeogr.).

## 2. Das syntaktische System.

In der syntaktischen Komponente einer Sprachbeschreibung werden die kombinatorischen Eigenschaften und Relationen der Ausdrücke der Sprache beschrieben. Es handelt sich dabei erstens darum, den Segmentationszusammenhang und die Distributionseigenschaften der Satzkonstituenten auf möglichst ökonomische Weise durch eine Anzahl generell formulierter Regeln herzuleiten; dies geschieht, indem zunächst Basis-Phrase-Marker (=Tiefenstrukturen) mittels geordneter kontextfreier Phrasenstrukturregeln, evtl. kontextsensitiver Subkategorisierungsregeln und Lexikonregeln hergestellt werden, und anschließend die Tiefenstrukturen sukzessive durch eine Folge von Transformationsregeln in wohlgeformte abgeleitete Phrase-Marker (=Oberflächenstrukturen) übergeführt werden. Zweitens handelt es sich darum, den eingeführten Konstituenten syntaktische Merkmale (oder Merkmalskomplexe) zuzuordnen, die mehrere Funktionen (und oft einige zugleich) erfüllen können :

- sie dienen als Kontextbedingung für die Anwendung von Subkategorisierungsregeln,
- sie gewährleisten die richtige Auswahl aus dem Lexikon der Sprache,
- sie dienen als Bedingung für die Anwendung von Transformationsregeln,
- sie dienen als Bedingung für die Anwendung von Flexionsregeln (dabei sind Flexionsregeln als spezielle und in der Ableitungsgeschichte eines Satzes sehr spät anzuwendende

Transformationsregeln zu verstehen),  
- sie dienen als Terme, die die Übersetzung in das semantische System der Sprache vorbereiten, d.h. es gibt für sie korrespondierende semantische Ausdrücke. (Wenn man die Forderung aufstellt, daß die Übersetzungsregeln einfache Substitutionsregeln sein sollen - mit relativ gering spezifiziertem Kontext - ist es sinnvoll, die syntaktischen Merkmale bereits auch im Hinblick auf die semantische Repräsentation zu formulieren.)

Nach dem Gesagten ist es deutlich, daß die Zeitadverbiale als Konstituenten aufgefaßt werden müssen, die durch Phrasenstrukturregeln eingeführt werden, während die Tempusmorpheme in Form von syntaktischen Merkmalen zu repräsentieren sind. Eine Komplizierung der Sachlage ergibt sich daraus, daß die Tempusmorpheme einen anderen Status besitzen als Numerus-, Genus- oder die meisten Personmorpheme: Erstens sind die letzteren aus Eigenschaften von Nominalphrasen NP abzuleiten, während die Tempusmorpheme nur dem Verb zukommen; nimmt man aber an, daß jeder Satz der Tiefenstruktur nur genau ein Verb enthält, kann man die Tempusmorpheme aus Eigenschaften des Satzes als Ganzes herleiten (etwa aus Relationen, die die im Satz ausgedrückten Sachverhalte betreffen). Zweitens sind die Numerus-, Genus- und Personmorpheme (abgesehen von denen der 1. und 2. Person) unabhängig von möglichen Sprech- oder Äußerungssituationen des Satzes, in dem sie vorkommen, zu interpretieren, hingegen sind die Tempusmorpheme in der Mehrzahl ihrer Verwendungen von deiktischem Charakter, d.h. sie müssen in Abhängigkeit von möglichen Sprechsituationen interpretiert werden. Dieselbe deiktische Eigenschaft wie die Tempusmorpheme besitzt auch eine Anzahl von Zeitadverbialen, z.B. neulich, morgen, in einer Woche, bisher. Die Verwendung solcher Ausdrücke impliziert gewisse Tempusmorpheme, z.B. ist mit neulich nur Präteritum, Perfekt oder Plusquamperfekt kompatibel, mit morgen nur Präsens oder Futur. Sätze wie  
x ich fand morgen die Lösung des Problems  
x ich werde gestern blau machen  
sind dagegen abweichend.

Die genannten Besonderheiten der Tempusmorpheme veranlassen mich zu einer gegenüber der üblichen Auffassung abweichenden Darstellung.

(1) Ich betrachte Tempus nicht als eine Kategorie, der in der Tiefenstruktur ein Konstituentenstatus zukommt (meistens wird Tempus als Subkonstituente einer Konstituenten Aux entwickelt), sondern lediglich als syntaktische Merkmalsinformation für die Flexionsbildung des Verbs.

(2) Tempus wird nicht der Verbalphrase VP zugeordnet, sondern der Satzkonstituenten (und wird in Form eines Tempusmorphems am Verb realisiert, wenn dem Satz das Merkmal !Finit! gegeben wird).





Es ist hinzuzufügen, daß die Substruktur der Zeitadverbiale noch etwas differenzierter formuliert werden muß, um allen empirischen Beobachtungen zu genügen. Insbesondere ist das Merkmal [e übl e<sub>k</sub>] noch in mehrere Teilmerkmale aufzuspalten. - Außerdem lassen sich mit dem hier wiedergegebenen Regelsystem noch nicht Ausdrücke wie heute, in diesen Tagen einführen, die zwar sprechzeitbezogen sind, aber dennoch mit allen Tempora kompatibel sind. - Das deutsche Perfekt ist erstens als Alternante zum Präteritum aufzufassen, zweitens wird es auch wie das englische Perfekt verwendet. Dieser zweite Gebrauch wird mit dem gegebenen Regelsystem nicht erfaßt. - Präpositionen, Konjunktionen und Adverbien (z.B. nach, nachdem, nachher) werden gemeinsam unter der Konstituenten adv<sub>k</sub> entwickelt. Sie werden erst durch nachfolgende Transformationsprozesse, die den Status der Ko-Konstituenten S<sub>k</sub> verändern, voneinander differenziert.

Ann.3) Mit dieser Maßnahme wird z.T. auf Vorschläge von Ross und Sadock zurückgegriffen.

John R. Ross, "On Declarative Sentences". Demnächst in R. Jacobs, P. Rosenbaum (eds), Readings in Transformational Grammar.

Jerold M. Sadock, "Hypersentences". Dissertation Univ. of Illinois. Urbana 1968.

4) Es ist möglich, eine noch etwas abstraktere Ebene der syntaktischen Repräsentation zu wählen. Auf ihr gibt es keine Konstituentenkategorie Advb, vielmehr wird bei Vorliegen einer bestimmten Geometrie des Phrase-Markers eine in ihm enthaltene VP als Advb interpretiert. Der Ausgangs-Phrase-Marker hat dann die Struktur  $[[[PRO - S]_{NP} - [V - [PRO - S]_{NP}]]_{VP}]_S$ .

In ihm werden gleichzeitig beide PRO-Elemente und die darüberstehenden NP-Knoten getilgt. Die Knoten des Transforms lassen sich entsprechend der hier gewählten Notation umbenennen:  $V \Rightarrow adv$ ,  $VP \Rightarrow Advb$ , so daß man  $[S' - [adv - S_k]_{Advb}]_S$  erhält.

### 3. Das semantische System.

Für die semantische Sprachbeschreibung wird ein normales prädikatenlogisches System zugrundegelegt, mit der folgenden deskriptiven Erweiterung:

Es wird eine Menge von Variablen t über Zeitintervallen z und eine primitive binäre Ordnungsrelation "vor" eingeführt.

Die Menge aller Zeitintervalle sei Z.

Die primitive Aussage "t<sub>1</sub> vor t<sub>2</sub>" besagt: wenn t<sub>1</sub> durch eine Konstante z<sub>1</sub> ersetzt wird und t<sub>2</sub> durch eine Konstante z<sub>2</sub>, dann geht z<sub>1</sub> ganz z<sub>2</sub> voraus.

(Allgemein gilt : verschiedene Variablen können nur durch verschiedene Konstanten, gleiche Variable müssen durch ein und dieselbe Konstante ersetzt werden.)

Die Eigenschaften von "vor" werden axiomatisch festgelegt.

Ax 1. Asymmetrie von vor.

$$t_1 \text{ vor } t_2 \supset \sim t_2 \text{ vor } t_1$$

Ax 2. Transitivität von vor.

$$t_1 \text{ vor } t_2 \wedge t_2 \text{ vor } t_3 \supset t_1 \text{ vor } t_3$$

Abgeleitete Relationen sind "nach" und "übl".

Def1.  $t_1$  nach  $t_2$  =df  $t_2$  vor  $t_1$

Def2.  $t_1$  übl  $t_2$  =df  $\sim t_1 \text{ vor } t_2 \wedge \sim t_2 \text{ vor } t_1$

Ax 3. Nichtverzweigung bzw. Linearität von Z.

$$(\langle t_1, t_2 \rangle) (t_1 \text{ vor } t_2 \vee t_1 \text{ nach } t_2 \vee t_1 \text{ übl } t_2)$$

Aus Ax1 und Ax2 folgt die Nichtzyklizität von Z im Endlichen.

Weiter lassen sich "teil" und "mom" (für Moment) definieren.

Def3.  $t_1$  teil  $t_2$  =df  $(t) (t \text{ übl } t_1 \supset t \text{ übl } t_2)$

Def4.  $\text{mom } t_1$  =df  $(t) (t \text{ teil } t_1 \supset t_1 \text{ teil } t)$

Ein Spezialfall der Überlappung ist die Gleichzeitigkeit (= "glz")

Def5.  $t_1$  glz  $t_2$  =df  $t_1 \text{ übl } t_2 \wedge (t) (t \text{ teil } t_1 \equiv t \text{ teil } t_2)$

Jede Variable  $t$  läßt sich als mögliches Argument einer prädikatenlogischen Formel verstehen. Formeln mit bis auf  $t$  konstanten Argumenten sind in der gewöhnlichen Prädikatenlogik Sätze ( bzw. Propositionen  $P$  ), sie repräsentieren, sofern sie die Wahrheitsbedingungen erfüllen, Variable  $V$  über mögliche Ereignisse ; bei Ersetzung von  $t$  durch eine Konstante  $z$  repräsentieren sie mögliche Ereignisse  $E$ . (Ereignisse brauchen durchaus nicht nur Sachverhalte zu sein, sie können auch fiktiv gedacht werden.) Mithilfe der Ordnungsrelation "vor" kann man zugleich auch eine Ordnung der Variablen über mögliche Ereignisse erreichen :

Ax 4.  $E P_1(\dots, t_1, \dots) \wedge E P_2(\dots, t_2, \dots) \wedge t_1 \text{ vor } t_2$

$$\wedge \text{wahr } P_1 \wedge \text{wahr } P_2 \supset V_1 \text{ vor } V_2$$

Für das Zeitsystem lassen sich mithilfe der Menge der reellen Zahlen  $R$  mehrere extensionale Modelle angeben.

(1) Zeitintervalle werden als Elemente der Potenzmenge  $P(R)$  repräsentiert, und zwar kommen alle zusammenhängende Teilmengen von  $R$  infrage.



(2) Zeitintervalle werden als reelle Zahlen  $x$  mit Umgebungen  $U(x)$  repräsentiert. (Für die Umgebungen sollen die Hausdorffschen Axiome gelten.)

(3) Approximativ werden alle Zeitintervalle, die Momente sind, als Zeitpunkte aufgefaßt. Dann sind alle nicht-momentanen Zeitintervalle als Intervalle von Zeitpunkten zu verstehen; sie lassen sich wie im Modell 1 durch Zahlenintervalle, d.h. durch zusammenhängende Teilmengen von  $\mathbb{R}$  repräsentieren.

Theoretisch macht das Modell 3 die geringsten Schwierigkeiten. Doch muß, um die Verhältnisse in den natürlichen Sprachen adäquat zu beschreiben, auch Modell 1 oder Modell 2 herangezogen werden. Dabei scheint Modell 2 die größeren Vorteile zu bieten:

(1) Ausdrücke wie jetzt oder heute beschreiben Umgebungen der Sprechzeit, die - je nach Kontext - klein oder groß sein können: ich bin jetzt müde vs. es ist jetzt Sommer vs. wir leben jetzt in einer Zwischeneiszeit.

(2) Ausdrücken wie gleich vs. bald, vorhin vs. neulich kann keine präzise Bedeutung gegeben werden. Auch die sie unterscheidende Distanz von der Sprechzeit kann nur vage, d.h. als Abstand zwischen zwei Umgebungen angegeben werden.

(3) Selbst Aussagen, die Ausdrücke des öffentlichen Zeitsystems verwenden, haben nicht immer präzise Bedeutung, sie können aber präzisiert werden:

das Unglück passierte um 9 Uhr → das Unglück passierte  
genau um 9 Uhr

vor 3 Jahren wurde meine Tochter geboren  
→ vor 3 Jahren, 4 Tagen und 2 Stunden wurde meine  
Tochter geboren.

(4) Mithilfe von Quantoren wie immer, stets wird eine Vergrößerung der Topologie erreicht, aber nur selten eine größte Topologie, d.h. die Gültigkeit der Aussage wird meistens nur für alle Zeiten innerhalb einer größeren Umgebung festgelegt, aber nicht für alle Zeiten überhaupt: z.B. du bist immer so nervös.

Unter den Variablen werden folgende Variablen ausgezeichnet:  
 $t_e$  über Ereigniszeitintervalle  
 $T_e$  über Mengen von Ereigniszeitintervallen  
 $t_s$  über Sprechzeitintervalle

Zur Übersetzung aus der syntaktischen Sprachbeschreibung in die semantische Sprachbeschreibung dienen die Referenzindizes  $e$  und  $s$  und die zwischen ihnen etablierten Merkmalsrelationen.

#### Übersetzungsregeln

1. (a) Ersetze jeden Index  $e$  durch eine Variable  $t_e$  oder  $T_e$ .
- (b) Ersetze jeden Index  $s$  durch eine Variable  $t_s$ .
- (c) Ersetze dabei gleiche Indizes durch dieselbe Variable und verschiedene Indizes durch verschiedene Variablen.

2. (a) Ersetze jedes Merkmal [e rel s] durch eine Menge von Paaren  $\{\langle T_e, t_s \rangle\}$  mit der Bedingung  $(\exists t)(t \in T_e \wedge t \text{ rel } t_s)$
- (b) Ersetze jedes Merkmal [e rel e<sub>k</sub>] durch eine Menge von Paaren  $\{\langle T_e, t_{ek} \rangle\}$  mit der Bedingung  $(\exists t)(t \in T_e \wedge t \text{ rel } t_{ek})$
- (c) Ersetze jede Merkmalkombination [Kont] [e rel s] durch eine Menge von Paaren  $\{\langle T_e, t_s \rangle\}$  mit der Bedingung  $(t)(t \text{ rel } t_s \supset t \in T_e)$
- (d) Ersetze jede Merkmalkombination [Kont] [e rel e<sub>k</sub>] durch eine Menge von Paaren  $\{\langle T_e, t_{ek} \rangle\}$  mit der Bedingung  $(t)(t \text{ rel } t_{ek} \supset t \in T_e)$

rel = vor, nach, übl

In (c) und (d) ist t evtl. auf eine größere Umgebung  $U(t_s)$  bzw.  $U(t_{ek})$  einzuschränken.

Bei einer Anzahl von Zeitadverbialen wie vor 4 Tagen, morgen, neulich, usw. ist die variable Menge  $T_e$ , aus der die Ereigniszeiten gewählt werden können, einzuschränken.

Z.B. gilt für morgen :  $T_e = \text{Tag}_{i+1}$ , wenn  $t_s \in \text{Tag}_i$ ,

und "Tag" ist entweder vage durch die normalen Wachperioden eines Sprechers oder präzise durch das Uhrensystem definiert.

#### 4. Anwendungsmöglichkeiten.

Das in den Abschnitten 2 und 3 erläuterte System der Zeitausdrücke und Zeitreferenzen ist gegenwärtig noch nicht genügend entwickelt worden, um bereits praktische Anwendungen zu ermöglichen. Deshalb kann ich nur skizzieren, in welcher Richtung Anwendungen möglich erscheinen bzw. geplant sind.

##### 4.1. Zeitalgorithmus.

Die Hauptaufgabe wird darin bestehen, einen Algorithmus zu entwickeln, der die einem Text inhärente Zeitstruktur auffinden kann, d.h. die zeitlichen Beziehungen zwischen den in den einzelnen Teilen des Textes beschriebenen Sachverhalten. Dieser Algorithmus muß an den möglichen Oberflächenstrukturen von Sätzen orientiert sein; er muß sie zurückführen auf entsprechende Tiefenstrukturen. Allerdings ist keine vollständige Rekonstruktion der Tiefenstrukturen notwendig, es genügt eine partielle Rekonstruktion, nämlich des zugrundeliegenden Systems von Ereigniszeitindizes, damit die Abbildung einer Textsequenz auf die ausgedrückten Ereignissequenzen möglich wird.

(1) Allen Ausdrücken, die - grob gesagt - auf Aktionen, Prozesse, oder Zustände verweisen, und daher tiefenstrukturell Satzstatus haben, werden Referenzindizes zugeordnet : primär allen finiten Verbkonstruktionen, ferner den Partizipial- und Infinitivkonstruktionen und den Nomina actionis, die als Nominalisierung von zugrundeliegenden Sätzen verstanden werden können, außerdem allen Uhrzeit- und Datumsangaben.

(2) Die Relationen zwischen den Referenzindizes werden bestimmt. Dafür sind zu analysieren

- die Satzsequenz des gegebenen Textes,
- alle Tempussequenzen (z.B. Präteritum-Plusquamperfekt),
- die Abhängigkeitsverhältnisse zwischen eingebetteten Sätzen und den Verben der dominierenden Sätze,
- alle vorkommenden zeitadverbialen Ausdrücke (Adverbien, Konjunktionen, Präpositionen),
- die Ausdrücke mit impliziter Zeitreferenz (z.B. Kausalkonjunktionen).

(3) Die Referenzindizes und die Relationen zwischen Referenzindizes werden in eine semantische Repräsentation übersetzt. Bei Vorkommen von Ausdrücken, die explizit auf externe Zeitskalen referieren (Uhrzeit- und Datumsangaben), werden die Variablen über Zeitintervalle ersetzt durch die entsprechenden Zeitkonstanten.

Für die Konstruktion eines derartigen Algorithmus müssen zunächst noch verschiedene Voraussetzungen erfüllt sein :

(1) Es muß eine vollständige Liste aller zeitreferierenden Ausdrücke, zusammen mit ihren syntaktischen Umgebungen bzw. ihrem syntaktischen Verhalten, aufgestellt werden. Wenn nur die am häufigsten vorkommenden Ausdrücke in der Liste verzeichnet werden sollen, muß eine mechanische Sichtung von Textproben vorgenommen werden.

(2) Es muß ein Teilalgorithmus entwickelt werden, der die Kompatibilität von Zeitausdrücken prüft, z.B. die Kompatibilität von Tempusformen mit bestimmten Klassen von Zeitadverbien. Der Prüfalgorithmus operiert über die semantischen Repräsentationen der Ausdrücke, z.B. mithilfe Durchschnittsbildungen. Bei leeren Durchschnitten besteht Nichtkompatibilität.

(3) Neben der zeitreferentiellen Struktur von Texten muß - wenigstens teilweise - die koreferentielle Struktur von Nominalphrasen NP einbezogen werden; vor allem müssen identische NP markiert und die Form von Artikeln und pronominalen Ausdrücken untersucht werden. Erst im Vergleich der zeitreferentiellen Struktur eines Textes mit seiner nominalen Verweisungsstruktur kann festgestellt werden, ob in dem untersuchten Text verschiedene voneinander unabhängige Prozesse oder nur Teilprozesse eines einzigen Prozesses dargestellt werden.

#### 4.2. Basisstrukturen.

Für die Klassifizierung (und damit auch eine übersichtliche Repräsentation) von zeitreferentiellen Beziehungen in Texten ist es nützlich, die möglichen Grundstrukturen zu untersuchen. Grundsätzlich wird angenommen, daß in Texten Sequenzen von Ereignissen oder von Teilprozessen E (bzw. Variable über Ereignissequenzen) dargestellt werden. (Auch stationäre Zustände werden als Teilprozesse verstanden.) Alle Sequenzen sind hinsichtlich der Zeitvariablen geordnet, d.h. mit der Sequenz  $(E_1, E_2, E_3, \dots, E_{n-1}, E_n, \dots)$  sind zugleich die Relationen  $t_1 \text{ vor } t_2, t_2 \text{ vor } t_3, \dots, t_{n-1} \text{ vor } t_n, \dots$  gegeben.

Folgende Ereignissequenzen sind möglich :

- lineare Sequenzen
  - a) einfache Sequenzen  $(E_1, E_2, \dots, E_n, \dots)$
  - b) parallele Sequenzen  $(\begin{Bmatrix} E_1 \\ E'_1 \end{Bmatrix}, \begin{Bmatrix} E_2 \\ E'_2 \end{Bmatrix}, \dots, \begin{Bmatrix} E_n \\ E'_n \end{Bmatrix}, \dots)$
- sich schneidende Sequenzen
  - a) divergente Sequenzen  $(E_1, \begin{Bmatrix} E_2 \\ E'_2 \end{Bmatrix}, \dots, \begin{Bmatrix} E_n \\ E'_n \end{Bmatrix}, \dots)$
  - b) konvergente Sequenzen  $(\begin{Bmatrix} E_1 \\ E'_1 \end{Bmatrix}, \begin{Bmatrix} E_2 \\ E'_2 \end{Bmatrix}, \dots, \begin{Bmatrix} E_{n-1} \\ E'_{n-1} \end{Bmatrix}, E_n, \dots)$

Jede dieser Sequenztypen ist in verschiedener Weise verbalisierbar, bzw. einer Folge von Sätzen zuzuordnen. (Das Ereignis  $E_i$  wird in einem Satz  $S_i$  verbalisiert. Daß einzelne der Sätze auch als eingebettete Sätze bzw. als Nebensätze erscheinen können, wird hier vernachlässigt.)

- einfache Ereignissequenzen :
  - Progression  $S_1 \wedge S_2 \wedge \dots \wedge S_n$
  - Regression  $S_n \wedge S_{n-1} \wedge \dots \wedge S_1$
  - Regression + Progression  $S_n \wedge S_1 \wedge S_2 \wedge \dots \wedge S_{n-1} (\wedge S_n)$
- parallele Ereignissequenzen :
  - $S_1 \wedge S'_1 \wedge S_2 \wedge S'_2 \wedge \dots \wedge S_n \wedge S'_n$
  - $S_1 \wedge S_2 \wedge \dots \wedge S_n \wedge S'_1 \wedge S'_2 \wedge \dots \wedge S'_n$
  - usw.
- sich schneidende Ereignissequenzen ( der Einfachheit halber wird  $n=2$  gesetzt)
  - divergente :
    - $S_1 \wedge S_2 \wedge S'_2$
    - $S_2 \wedge S'_2 \wedge S_1$
    - $S_2 \wedge S_1 \wedge S'_2$

konvergente :  $S_1 \wedge S'_1 \wedge S_2$   
 $S_2 \wedge S_1 \wedge S'_1$   
 $S_1 \wedge S_2 \wedge S'_1$   
 usw.

Zur Illustration der Verbalisierung von sich schneidenden Ereignissequenzen mögen z.B. folgende Satzfolgen dienen:  
 Hans und Fritz trafen sich am Theater. ^ Nachher spielte Hans Fußball und Fritz lernte Latein.  
 Hans spielte Fußball. ^ Fritz lernte Latein. ^ Vorher hatten sie sich am Theater getroffen.  
 Hans spielte Fußball. ^ Vorher hatte er Fritz, der später Latein lernte, am Theater getroffen.  
 Hans spielte Fußball. ^ Fritz lernte Latein. ^ Nachher trafen sie sich am Theater.  
 Hans und Fritz trafen sich am Theater. ^ Vorher hatte Hans Fußball gespielt, während Fritz Latein lernte.  
 Hans spielte Fußball. ^ Nachher traf er am Theater Fritz, der Latein gelernt hatte.

Nach diesen Vorbereitungen läßt sich allgemein die Zeitstruktur eines Textes aus  $n$  Sätzen als geordnetes Quadrupel

$\langle \text{UTV}, \text{EOR}, \text{Id}, \text{R} \rangle$  einführen, mit  
 UTV für unmittelbaren Textvorgänger  
 EOR für Ereignisordnungsrelation  
 Id für Identitätsrelation  
 R für Abbildung aus der Menge der Sätze in die Menge der Ereignisse

und den Bedingungen

UTV  $\subset$  strikte Ordnung  
 EOR  $\subset$  Ordnung  
 Kardinalzahl von EOR =  $n$   
 Id  $\subset$  EOR  
 Feld(R) = Feld(UTV)  $\cup$  Feld(EOR)  
 Vorbereich(R) = Feld(UTV).

Wenn innerhalb des Textes nicht auf andere Sätze derselben Struktur bezuggenommen werden soll, sind Feld(UTV) und Feld(EOR) distinkt.

#### 4.3. Anwendungsgebiete.

Ich sehe die folgenden Gebiete, auf den eine mechanische Gewinnung, Speicherung oder Verbalisierung von Zeitstrukturen Vorteile verspricht.

(1) Textaufbereitung für Information Retrieval.

Charakterisierung von Texten, Berichten usw. nach den zugrundeliegenden Beziehungen zwischen Satzsequenzen und ausgedrückten Ereignissequenzen. Diese Charakterisierung kann mithilfe der in 4.2 angedeuteten Basisstrukturen erfolgen.

(2) Auswertung von psycho- und soziolinguistischen Experimenten.

Ein solches Experiment kann darin bestehen, daß der Versuchsleiter verschiedene Texte vorliest, die dieselben Ereignissequenzen ausdrücken, sich aber hinsichtlich der verbalen Organisation unterscheiden. Die Versuchspersonen werden jeweils angehalten, den Inhalt eines vorgelesenen Textes frei wiederzugeben. Aus den Fehlern oder Umstellungen bei der Textwiedergabe kann auf psychologische Mechanismen der Textwahrnehmung und des Texterinnerns geschlossen werden, wie auch auf evtl. schichtenspezifische Unterschiede in der Ausbildung dieser Mechanismen.

(3) Rekonstruktion und Speicherung von historischen oder anderen zurückliegenden Ereigniszusammenhängen in Form von Zeitflußdiagrammen. Die zu analysierenden Texte sind z.B. historische Dokumente oder sind Zeugenaussagen in einem juristischen Prozeß. Anhand von Zeitflußdiagrammen kann leicht die Konsistenz der historischen Daten oder der Zeugenaussagen geprüft werden.