

Using Discourse Focus, Temporal Focus, and Spatial Focus to Generate Multisentential Text

Mark T. Maybury

Rome Air Development Center
Intelligent Interface Group (COES)
Griffiss AFB, Rome New York 13441-5700
maybury@tops20.radc.af.mil

and

Computer Laboratory
Cambridge University, England CB2 3QG

Abstract

This paper claims that reliance on discourse focus to guide the production of rhetorically structured texts is insufficient over lengthier stretches of prose. Instead, this paper argues that at least three distinct attentional constraints are required: discourse focus [Sidner, 1979, 1983; Grosz and Sidner, 1986], temporal focus [Webber, 1988], and a novel notion of spatial focus. The paper illustrates the operation of this tripartite theory of focus in a computational system (TEXPLAN) that plans multisentential text.

Introduction

Effective generation of prose demands not only knowledge of rhetorical structure but also rich models of entities, events and states, knowledge of tense and aspect, and mechanisms to track focus of attention with respect to discourse, time, and space. McKeown [1982] used discourse focus (DF) [Sidner, 1979, 1983] to guide the selection, order, and realization of rhetorical schema-based descriptions of database contents. McKeown suggested the following focus shift preferences to mediate among competing propositional content:

1. shift DF to an entity mentioned in the previous proposition
2. maintain current DF
3. resume a past DF
4. shift DF to an entity most related to the current DF

The three global registers (past, current, and potential focus) tracked DF and were updated by examining the content of a rhetorical proposition (instantiated with information from a knowledge base) guided by the type of the rhetorical predicate (e.g., identification, attributive).

In contrast to schema-based systems, recent work based on Rhetorical Structure Theory (RST) [Mann and Thompson, 1987] attempts to produce effective text using plan-based strategies. Only Hovy's [1988] implementation of RST has examined the task of conveying events and states. Hovy's [1988] "structurer" uses his *sequence* RST operator to produce the following

narration of events in a naval domain (where C4 indicates a condition or level of operational readiness):

*Knox, which is C4, is en route to Sasebo.
Knox, which is at 18N 79E, heads SSW. It
arrives on 4/24. It loads for 4 days.*

To produce this text, Hovy's *sequence* operator is given a beginning "action." The nucleus of the *sequence* operator allows the text to "grow" and indicate the circumstances, attributes, and/or purpose of this action. Similarly, the satellite of the *sequence* operator allows the text to indicate the attributes and/or details of the next contiguous action in some sequence. The satellite also includes a recursive call to the *sequence* operator for the next action.

Unfortunately, Hovy's operators, like text schema, fail to indicate what effects these orderings or the addition of information at growth points have on the hearer. Therefore, they fail to characterize the motivation for selecting among the different arrangements that narrative employs to achieve specific effects on the hearer (e.g., creating interest, suspense, or mystery). In addition, the plan operator does not consider states as first order objects in some causal chain (the fact "Knox is C4" is just an attribute extending off the "en route" event). This is important because states have complex relations (e.g., enablement, causation) to other states and events in the world. Finally, sequences are assumed to be contiguous and yet events are often simultaneous or overlapping in time [Allen, 1984].

This purely RST-based approach was improved upon by Hovy and McCoy [1989] by incorporating Focus Trees [McCoy and Cheng, 1988] to guide the ordering and interrelationships of sentence topics. This combined approach produced:

*With readiness C4, Knox is en route
to Sasebo. It is at 79N 18E heading
SSW. It will arrive 4/24 and will
load for four days.*

Text coherence is improved not only by regrouping content (a result of restrictions on the traversal of the Focus Tree) but also by using tensed verbs (e.g., future tense of "arrive" in the last utterance) to explicitly indicate

the temporal relations among events. Unfortunately, no details of how this tense is generated are provided. Furthermore, examination of human generated prose indicates that not only DF but additional constraints on temporal focus and spatial focus are necessary to produce lengthier prose.

Therefore, the remainder of this paper first details a tripartite theory of focus and proposes focus shift rules. Next an ontology of events and states is introduced that serves as the basis for a model of tense and aspect which, guided by temporal focus, is used to verbalize events and states. The temporal organization and realization of events is exemplified in the context of report generation. Finally, an example from a route planner is given to illustrate spatial organization and the use of spatial focus.

Discourse Focus, Temporal Focus, and Spatial Focus

Like Hovy [1988], TEXPLAN uses a hierarchical planner to select, structure, and order propositional content using a library of plan operators (detailed in a subsequent section). To achieve a given discourse goal (e.g., get the reader to know about an event), the planner selects among competing plan operators using general plan operator selection heuristics [Moore, 1989] such as prefer plan operators that meet all preconditions, that have fewer subplans, that have fewer new variables and so on. The leaf nodes of the resulting text plan are speech acts with associated propositional content in the form of rhetorical predicates [c.f. McKeown, 1982]. As in McKeown's TEXT, TEXPLAN tracks past, current, and potential discourse focus (DF) in global registers. When the text planner selects a particular rhetorical proposition, the attentional mechanism extracts the default discourse focus (a position associated with each rhetorical predicate) and updates the global registers. This focus information is then used to guide surface choice.

In contrast to DF, Webber [1988] proposed Temporal Focus (TF) as the event currently being focused on temporally and suggests that TF is used to integrate events into some evolving spatio-temporal event/situation structure. TF can shift depending on the relations that hold between events and their times of occurrence. Webber [1988] suggests three TF shifts: maintenance, forward, and backward. Nakhimovsky [1988] classifies local TF shifts as: forward, sideways, and backward "micromoves". Forward and backward shifts correspond to introducing the consequence or preparatory phases of events [Moens and Steedman, 1988]. Backward shifts start a new discourse segment. In TEXPLAN, TF indicates the Reichenbachian [1947] reference time. TF shifts (local or micromoves) are implemented via the plan operators and are ordered as follows:

1. Maintain current TF (maintenance)
2. TF progresses "naturally" forward (progression)
3. Shift TF to a simultaneous event/state (lateral shift)

In addition two other long distance temporal shifts are possible but are not addressed in the current implementation:

4. Shift TF to a prior event/state. (flashback)
5. Shift TF to a distant future event/state. (flashforward)

Temporal shifts are conveyed to the reader in part by verb tense and aspect as in the use of future tense in "John just arrived. He was in an accident yesterday and ..." Temporal shifts are also indicated by adverbials (e.g., "five minutes later"), explicit references to time ("at seven p.m."), and cue words (e.g., "simultaneously"). TEXPLAN tracks TF by recording pointers to events that appear in the propositional content selected by the text planner just as it records DF from selected propositional content following McKeown (1982). As with DF, past, current, and potential temporal focus registers are updated after each utterance.

Just as discourse can be topically and temporally organized, psychologists have observed that humans utilize spatial organizations, for example when people describe their apartments [Linde and Labov, 1975]. Shifts analogous to those of DF and TF can occur along the dimension not of discourse or time but rather *space*. I define spatial focus (SF) as the current entity or group of entities (and its/their associated spatial location) that the reader is attending to in space. The notion of spatial focus is related but distinct from Conklin's (1983) notion of *visual saliency*. Visual saliency is the noteworthiness (from one perspective) of an entity in relation to a set of static objects. Spatial focus, in contrast, refers to a currently focused entity (a "moving target") that is spatially related to the other entities currently in the background (static entities) or foreground (dynamic entities). Just as DF and TF follow regular shifts, the following ordered legal shifts appear to govern SF:

1. Maintain the current SF
2. Shift SF to an entity spatially related to the current SF
3. Shift SF to some distant point or region.

Shifts in rule 2 can be relational (e.g., behind, in-front-of, left-of, right-of, above, below, on-top-of, etc.) or in terms of distance (e.g., "five miles away"). Shifts in rule 3 signal a new discourse segment. Just as TF can refer to points or intervals of time, SF can refer to either a point in space ("At 23° latitude 5° longitude"), a region (e.g., "In Chesterville today, ...") or a set of points or regions (analogous to discourse focus spaces [Grosz, 1977]). After each utterance, by examining the underlying propositional content TEXPLAN updates global registers that encode the past, current, and potential spatial foci. In the current implementation, the system prefers topical over causal over temporal over spatial orderings.

The next sections illustrate the input to the generator, and how TEXPLAN uses the notions of Reichenbachian time and temporal focus to narrate events and states. By

tracking TF and exploiting the temporal information in the underlying event/state model, TEXPLAN is able to select, order, and linguistically realize events and states. The realization component of the system selects proper verb tense and aspect and indicates shifts in TF, for example, through the use of adverbials. A final section illustrates the use of SF in locative instruction (i.e., route plans).

Event and State Ontology

As Hovy and McCoy's example in the introduction illustrates, more sophisticated representations of verb tense and aspect is key to generating coherent narrative text. This demands a more sophisticated representation of events, states, and their relationship to tense and aspect. Representing and linguistically realizing events concerns issues of temporality, causality, and enablement as well as verb tense and aspect. Discussion of noninstantaneous events dates at least to Aristotle's distinction between process (*energia*) and state (*stasis*) and these issues have been the focus of attention in philosophy, linguistics, and computational linguistics [c.f. Allen, 1988]. While an ontology of events and states is beyond the scope of this paper, it is necessary as a starting point for generation to indicate the nature of the underlying propositional content and so we make a few intuitive distinctions.

Events are physical, linguistic, or psychological happenings at some time and place. States, in contrast, refer to perpetual or temporally unbounded conditions such as the physical, psychological, or emotional state of an agent or entity. States include relations that hold between agents or entities (e.g., possession, ownership) [Nakhimovsky, 1988]. This classification is but one (conceptual) classification of events and states. Ehrich [1986], for example, uses the features of duration, resultativity, and intentionality to produce an orthogonal categorization.

Processes, in contrast to states, involve changes or transformations over the interval for which they hold and often have some associated rate of progress toward a goal or a rate of consumption of resources [Nakhimovsky, 1988]. Nakhimovsky makes a key distinction between events, processes and states:

For a linguist, the distinction between **event-process** is one of aspectual perspective: "The term 'process' means a dynamic situation viewed imperfectly, and the term 'event' means a dynamic situation viewed perfectly" (Comrie, 1976: 51). The distinction **process-state** is one of aspectual class.

In this paper the term *event* is used to refer both to an instantaneous event (e.g., snap, click, wink) as well as events with a duration which can be viewed perfectly (event) or imperfectly (process).

A collection of related events and states constitutes an event/state network analogous to Webber's [1987] event/situation structure. This network of events and states serves as the basis for generation in TEXPLAN. In

the input to the text generator, each event or state is represented in a frame-like structure. Events and states have associated attributes, roles, and relationships. The term *attributes* refers to characteristics local to the event or state such as its time of occurrence (a point or interval), its type (e.g., physical, linguistic), and any constituents (i.e., subevents or substates). *Roles* refer to the semantic role an entity plays in the event or state (e.g., agent, patient). Finally, *relations* refer to the associated enablement(s), cause(s), and effect(s) of an event or state.

Tense and Aspect

The rich notions of time associated with events and states are conveyed in part through verb tense. English verb tense (e.g., simple past, present, and future; and past, present, and future perfect) relies on a tripartite notion of time which includes: the point or time at which the utterance is spoken (S), the point at which the event happens (E), and the point of reference (R) [Reichenbach, 1947]. R is the time "talked about" or "focused on" and in TEXPLAN corresponds to the above notion of TF. Because the absolute time of the event (E) appears in the event structure, the linguistic realization component can select the appropriate verb tense by reasoning about the time the speaker is narrating (S) (e.g., "now") and the time the overall narration focuses on (R). This contrasts with verb choice based solely on the underlying event structure [e.g., Kalita, 1989, p. 410]. Table 1 relates E, R, and S to tense where "<" indicates temporal antecedence and "=" indicates temporal simultaneity.

Time	Tense	Example
E=R=S	simple present	"John eats."
E=R<S	simple past	"John ate the beans."
S<E=R	simple future	"John will eat the beans."
E<R=S	present perfect	"John has eaten."
E<R<S	past perfect	"John had eaten the beans."
S<E<R	future perfect	"John will have eaten."

Table 1

This point-based time representation could be extended to consider time intervals [Allen, 1984].

TEXPLAN's sentence generator uses an admittedly simplified¹ prototypical verb sequence following Winograd [1983] (e.g., Modal + Have + Be1 + Be2 + Main-verb). Individual verbs include both *modals* such as "will", "can", "could" (which have only one form), and ordinary verbs which have five basic forms in third person, singular: infinitive (e.g., "to walk"), simple present ("walks"), simple past ("walked"), present participle ("walking"), and past participle ("walked"). Future tense does not have its own syntactic form and is implemented by the modals "will" or "shall".

In contrast to tense, aspect is a grammatical category of the verb implemented by affixes, auxiliaries, and so on

¹Matthiesen (1984) has discussed more general tense assignment.

[Nakhimovsky, 1988]. This arises from the temporal characteristics of the underlying event (i.e., point versus interval), the relationship of the reference time (R) to event time (E) (e.g., $E < R$ indicates perfective; $E = R$ imperfective) and the lexical aspect of the verb (e.g., the progressive "I am eating" versus the culmination "I am finishing"). The current implementation addresses only the first two cases. Perfective sequences use Have (e.g., has taken), progressive use Be1 (e.g., was taking), and passive use Be2 (e.g., was taken).

The LACE Simulation

With the representation of events and states, their temporal structure, and their relation to tense and aspect detailed, this section turns to the task of planning and realizing narrative. While text analysis motivated focus models and narrative plans (described below), there was a practical need to narrate simulation events and states in complex multi-agent simulations. In particular, narrative plans and temporal focus were tested in LACE (Land Air Combat in ERIC), a knowledge-based battle simulation system [Anken, 1989]. LACE is coded in ERIC, an object-oriented simulation language [Hilton, 1987]. Narration in LACE is complex because multiple, autonomous agents interact simultaneously to achieve their individual goals. For example, attacking forces attempt to bomb targets, refuel aircraft, move cargo, and suppress ground forces with electronic countermeasures. In contrast, defending forces attempt to detect, track, and destroy intruders. To give a feel for the nature of the sophistication of the simulation, there are over 150 classes each with dozens of behaviors. In a typical run of the simulation, hundreds (and potentially thousands) of instances of objects are generated. If several agents (e.g., 10 or 15) are given goals to pursue at the start of a simulation run, their actions generate thousands of events per minute as agents react to both the environment and to the behavior of other agents in the simulation. For example, if a long-range radar detects an intruding aircraft it will order its associated mobile surface-to-air-missile sites to electronically track, pursue (i.e., along the ground) and fire at the incoming target. The generation task, then, is to produce a report of the events after simulating conflicts between two opposing military forces. Over fifty texts were produced, an example of which is detailed below.

The input to narration generation is a network of events and states from the underlying simulation. Each machine second that the simulation clock ticks LACE records the situations that occur at that moment. The simulation measures time using Common Lisp's universal time (i.e., as seconds since the year 1900). These event snapshots (e.g., at time 34300023 #<search-radar-291> began sweeping) are slotted into the representation of events which includes their associated properties (i.e., attributes such as location and duration; relations to other events such as causal and temporal connections; and any associated roles such as the agent,

patient, and so on). Collectively these structures characterize the event/state network that represents an overall spatio-temporal-causal picture of the simulation. This event/state network is preprocessed to prune details. For example, persistent or uninteresting (e.g., frequent, non-unique, or unimportant) events can be deleted. Accessing this event/state network, TEXPLAN's narrative plans select, order, and realize events to compose a report.

Event selection from the event/state network is guided by the saliency of the occurrence. In the current implementation, event or state saliency is a function of:

1. the kind and amount of links associated with an event or state in the event/state network
2. the frequency of occurrence in the event/state network
3. domain-specific knowledge of importance

The first item concerns issues such as does the event achieve a main goal of a key agent in the simulation, does it motivate, enable, or cause a number of events or states to occur, and so on. The second item is simply the observation that frequent or commonplace events are boring. For example in LACE long-range radar are constantly sweeping, SAM¹ sites are always repositioning themselves, and aircraft are always flying point-based ingress/egress routes. An example of the third item is that mission types have an order of interestingness (e.g., offensive air attack > SAM suppression > refueling > transportation). This is analogous to Kittredge's et al. [1986] weather reports which indicate warnings first and then WINDS > CLOUD-COVER > PRECIPITATION > FOG&MIST > VISIBILITY. There are other issues involved in saliency that are beyond the scope of this paper such as the inferrability of events and states, event and state persistence, and the representation of perceptual saliency which requires complex user modelling. The next section addresses the key problem: how do we select, order, and present events from the event/state structure in a report?

Report Generation

The most basic form of narration recounts events in their temporal order of occurrence. This occurs in a journal, record, account, or chronicle, collectively termed a *report*. Reports typically consist of the most important or salient events in some domain during one period of time (e.g., stock market report, weather report, news report, battle report). Sometimes reports focus on events and states involving one dimension of an agent as in a medical record, an educational record, or a political record.

TEXPLAN plans narrative by reasoning about what effects certain rhetorical strategies or speech acts will have on the user if they are employed. This is accomplished by representing each *communicative act* (either a rhetorical act or a speech act) as an operator in a library of plans which are reasoned about by a hierarchical text planner [Sacerdoti, 1977] similar to that used by Hovy [1988] and

¹Surface to Air Missile

[Moore, 1989]. Communicative acts have specific constraints, enabling conditions, effects on the hearer, and decompositions. A *rhetorical act* characterizes the communicative function of one or more utterances (e.g., describe, define, compare, narrate) and may employ other rhetorical acts and/or speech acts to achieve its associated goals. In contrast, a *speech act* [Searle, 1969, 1975] refers to the illocutionary force of utterances (e.g., inform, request, warn, promise). The propositional content of a speech act is a *rhetorical predicate* whose function is to abstract particular kinds of information from a knowledge base (e.g., constituency predicates refer for subparts of entities whereas classification predicates refer to subtypes of entities, logical-definition predicates include the genus and differentia of an entity.) Over twenty rhetorical predicates and fifty plan operators have been implemented that are able to produce a variety of texts including description, narration, exposition, and argument.

Plan operators are represented in an extension of first order predicate calculus that allows for optionality in the decomposition. Like conventional planners, each plan operator defines the *constraints* and *preconditions* that must hold before a communicative act applies, its intended *effect*, as well as its refinement or *decomposition* into subacts. Constraints, unlike preconditions, cannot be achieved or planned for if they are false. In plan operators, variables are italicized (e.g., *H*, *S*, and *entity*) and constants appear in upper-case plain type. Intensional operators, such as WANT, KNOW and BELIEVE appear in capitals. KNOW details an agent's specific knowledge of the truth-values of propositions (e.g., KNOW(*H*, Red(ROBIN-1)) OR KNOW(*H*, ¬Yellow(ROBIN-1))) where truth or falsity is defined by the propositions in the knowledge base. That is, KNOW(*H*, *P*) implies $P \wedge BELIEVE(H, P)$. Of course an agent can hold an invalid belief (e.g., BELIEVE(JOHN, Yellow(ROBIN-1))). KNOW-ABOUT is a predicate that is an abstraction of a set of epistemic attitudes of some agent toward an individual. An agent can KNOW-ABOUT an entity or event (e.g., KNOW-ABOUT(*H*, ROBIN-1) OR KNOW-ABOUT(*H*, EXPLOSION-445)) if they KNOW its characteristics, components, subtypes, or purpose.

For example the top-level narration plan shown in Figure 1 encodes the communicative act of speaker (*S*) narrating some sequence of salient events in topical order so that the hearer (*H*) knows about them. Similar top-level plan operators narrate events causally, temporally and spatially. If the events can be sequenced in several ways, the planner prefers topical to causal to temporal to spatial orderings. The narrate-report-topically plan operator is chosen when there is no obvious temporal or spatial sequencing, for example, when multiple events occur simultaneously or when they occur in similar spatial locations.

After an event/state structure is captured from a typical run of the LACE simulation (e.g., some blue forces are attacking some red targets), the generation of a report is initiated by posting the top-level goal "narrate all the events in the event/state network". This matches the header of the plan operator in Figure 1 (as well as others) and the unordered list of events is bound to the *events* parameter of the header. This narrate-report-topically plan operator is selected because (1) relative to the number of events, there are few principle agents (i.e., missions) which enables topical grouping and (2) because other plan operators are less appealing (for example, a top-level temporal organization would be confusing because they are many simultaneous events involving different agents). The decomposition of the plan operator first introduces the events using the introduce plan operator which indicates the static background or framework within which the events in the foreground are to be interpreted. The introduce plan operator describes the principle time, place, agents (i.e., characters) using a variety of rhetorical means including definition, attribution, illustration, and division (i.e., classification and constituency) (see Maybury [1990] for details). In LACE information in the introduction is retrieved from the overall mission package which drives the entire simulation (represented in a frame-like structure). This package includes the time of the major missions, their location, their type, and so on. In this case this information is described using two rhetorical predicates: logical definition (which indicates the genus and differentia of the package) and constituency (which indicates the subparts, in this case missions).

NAME	narrate-report-topically
HEADER	Narrate(<i>S</i> , <i>H</i> , <i>events</i>)
CONSTRAINTS	Topical-Sequence(<i>events</i>) $\wedge \forall e \in events \text{ Event}(e)$
PRECONDITIONS	$\forall e \in events \text{ KNOW-ABOUT}(S, e)$
EFFECTS	$\forall t \in Topics(events) \text{ KNOW-ABOUT}(H, t)$
DECOMPOSITION	Introduce(<i>S</i> , <i>H</i> , <i>events</i>) \wedge $\forall topic \in Order-According-to-Salience(Topics(events))$ Narrate-Sequence(<i>S</i> , <i>H</i> , Events-with-Topic(<i>events</i> , <i>topic</i>))

Figure 1. Top-level, Uninstantiated Text Plan Operator for Report Narration

NAME	narrate-temporal-sequence
HEADER	Narrate-Sequence(<i>S</i> , <i>H</i> , <i>events</i>)
CONSTRAINTS	Temporal-Sequence(<i>events</i>) $\wedge \forall e \in \text{events} \text{Event}(e)$
PRECONDITIONS	$\forall e \in \text{events} \text{KNOW-ABOUT}(S, e)$
EFFECTS	$\forall e \in \text{events} \text{KNOW-ABOUT}(H, e)$
DECOMPOSITION	$\forall e \in \text{Select-and-order-chronologically}(\text{events})$ Inform(<i>S</i> , <i>H</i> , Event(<i>e</i>))

Figure 2. Uninstantiated Operator for Temporal Sequence Narration

Next the subjects or topics which the events concern are ordered according to saliency. Saliency is determined by the frequency, uniqueness, importance, and so on of events in the domain as detailed in the previous section. After events are grouped topically in order of saliency, the *narrate-temporal-sequence* plan shown in Figure 2 selects salient events and then exploits their temporal order to sequence them.

As plan operator decompositions are achieved, they are recorded in a hierarchical *text plan* (a communicative action decomposition) which records the structure, order, and content of the final text (see Figure 3). Each node in the text plan corresponds to a plan operator which formalizes a communicative act. The leaf nodes of the text structure indicate speech acts with associated *rhetorical propositions*, rhetorical predicates instantiated with propositional content from the event/state structure.

In Figure 3, the event sequence is introduced by describing the mission package which initiated the previous run of the simulation (#<air-strike-10>) by informing the hearer of its logical definition (i.e., superclass and distinguishing features) and then by indicating its constituents or components. Next events are narrated in topical groups, each of which is temporally sequenced. The text plan is linearized and each speech act with its rhetorical proposition is linguistically realized using a unification-based surface generator [Maybury, 1989]. The final surface form of the text structure in Figure 3 is shown in Figure 4. The surface generator includes an orthographic layout module that examines the text structure and begins a new paragraph when a new discourse segment begins (e.g., before the introduction and after each topical narration sequence).

While the plan operators guide the ordering of propositional content (and thus focus shifts), global registers for DF, TF, and SF are updated by examining the propositional content selected by the text planner. These

registers are used by the surface generator to guide surface choice. For example, a noun phrase is pronominalized if it is given (versus new) and was the previous current DF. And by examining the relationship of speaker time (*S*), event time (*E*), and reference time (*R* based on TF), the realization component is able to choose appropriate verb tense (e.g., if $E=R<S$ then use simple past). Similarly, the realization of temporal adverbs, which help to convey temporal relations among events, (e.g., "and then", "three minutes later", "simultaneously", "before", "after") is guided by the relationship of past to current TF. For example if the time of the current TF equals that of the past TF then connectives such as "simultaneously" or "at the same time" introduce utterances. Both temporal and spatial adverbs (exemplified in the next section) often refer anaphorically to the current TF (e.g., "ten minutes later") or SF (e.g., "three miles away"). During realization, events that are of the same class, occur at the same time and share a semantic patient are combined as in "One minute later Erfurt-A and Erfurt-D simultaneously fired a missile at offensive counter air mission 102." The adverb "again" is used when an event has already occurred, and thus functions as an anaphor.

Other classes of adverbs can also enrich the event description. These include adverbs regarding manner (e.g., "defly", "sadly"), rate (e.g., "slowly", "rapidly"), duration (e.g., "for twenty minutes"), frequency ("every ten minutes"), and numeration (e.g., "seventeen times"). Some adverbs (e.g., manner, rate, durative, locative) regard *internal* properties of the event whereas others relate the current event to other events and therefore are *external* such as temporal adverbs (e.g., "simultaneously", "yesterday") or "anaphoric" adverbs (e.g., "again", "as before"). The relationship of these and other classes of adverbials to DF, TF, and SF remains an interesting area for further work.

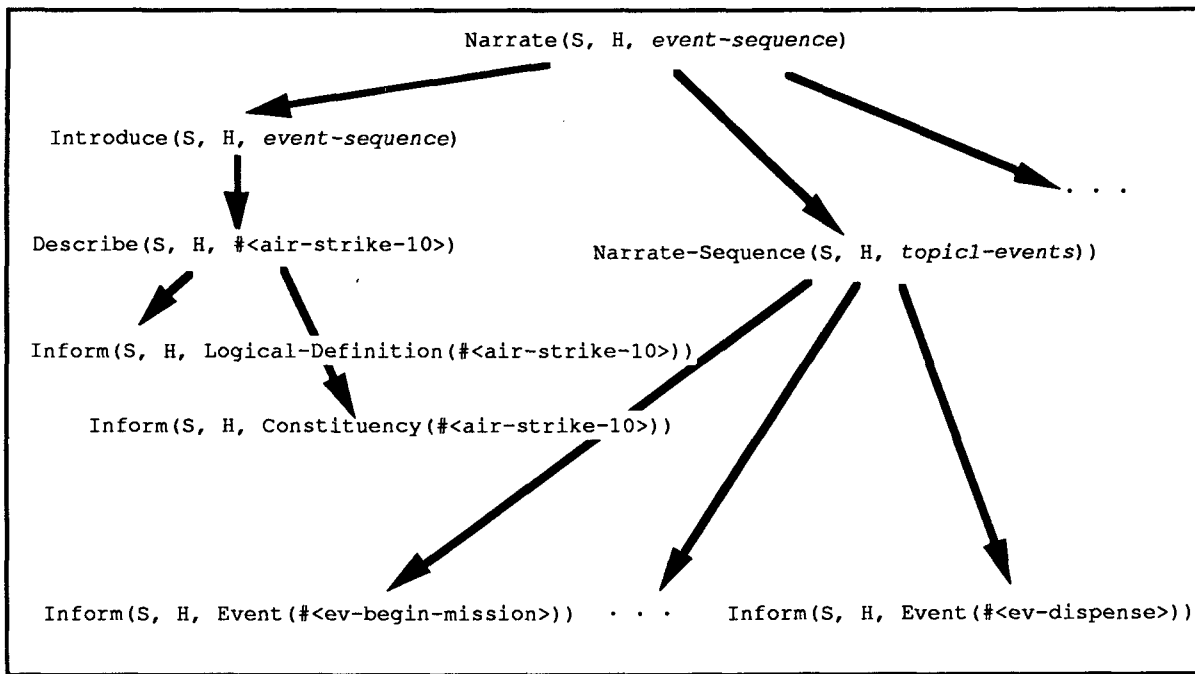


Figure 3. Narrative Text Structure of LACE report

Air-strike 10 was an attack against Dresden airfield in the Fulda-Gap region of West Germany on Tuesday December 2, 1987. Air-strike 10 included three Offensive Counter Air Missions (OCA100, OCA101, and OCA102), one SAM Suppression Mission (SSM444), one Transportation Mission (TRANS100), and one air refueling mission (RFL100).

Offensive Counter Air Mission 100 began mission execution at 8:20::0 Tuesday December 2, 1987. 902TFW-F-16c dispensed four aircraft for Offensive Counter Air Mission 100. Eight minutes later Offensive Counter Air Mission 100 began flying its ingress route. Three minutes later Allstedt-B and Allstedt-C simultaneously fired a missile at Offensive Counter Air Mission 100. And fifty-nine seconds later Offensive Counter Air Mission 100 was ordered to abort its mission. One second later Allstedt-C and Allstedt-B again simultaneously fired a missile at Offensive Counter Air Mission 100. Two minutes later Allstedt-B again fired a missile at Offensive Counter Air Mission 100. Then one minute later Erfurt-A fired a missile at Offensive Counter Air Mission 100. Then two minutes later Haina-B fired a missile at Offensive Counter Air Mission 100. Seven minutes later Offensive Counter Air Mission 100 ended its mission. It generated its post-mission report.

In the meantime SAM Suppression Mission 444 began mission execution at 8:30::0 Tuesday December 2, 1987. 126TFW-F-4g dispensed one aircraft for SAM Suppression Mission 444. SAM Suppression Mission 444 began flying its ingress route. Thirteen minutes later Mobile-SAM1 fired a missile at SAM Suppression Mission 444. Then fifty-nine seconds later SAM Suppression Mission 444 was ordered to abort its mission. And then one second later Mobile-SAM2 fired a missile at SAM Suppression Mission 444. One minute later Mobile-SAM2 and Mobile-SAM1 simultaneously fired a missile at SAM Suppression Mission 444.

In the meantime Offensive Counter Air Mission 101 began mission execution at 8:41::40 Tuesday December 2, 1987. 900TFW-F-4c dispensed four aircraft for Offensive Counter Air Mission 101. Then seven minutes later Offensive Counter Air Mission 101 began flying its ingress route. Then ten minutes later it bombed its target. It began flying its egress route. Thirty-Six minutes later it ended its mission. It generated its post-mission report.

Meanwhile Transportation Mission 250 ...

Figure 4. Temporally and Topically Focused LACE report

Locative Instructions

While space limitations prohibit a full explication of spatial plan operators, a short indicative example is provided. TEXPLAN produces locative instructions, that is directions that enable an agent to get from point a to point b, in the context of the Map Display System (Hilton, 1987; Hilton and Anken, 1990), a cartographic database underlying LACE which includes thousands of entities such as roads, bridges, airports and so on. For example if the user asks how to get from Mannheim to Heidelberg, the underlying application plans a route which TEXPLAN organizes and presents as:

From Mannheim take Route 38 Southeast for four kilometers to the intersection of Route 38 and Autobahn A5. From the intersection of Route 38 and Autobahn A5 take Autobahn A5 Southeast for seven kilometers to Heidelberg. Heidelberg is located in block 32umv7070 at 49.39° latitude and 6.68° longitude, 4 kilometers Northwest of Dossenheim, six kilometers Northwest of Edingen, and five kilometers Southwest of Eppelheim.

The structure of the text follows the spatial progression of the route. TEXPLAN tracks the spatial focus (each segment of the route) as the planner sequentially informs the user of each segment. This information is then used to guide the realization of locative adverbials and prepositions. For example, directionals such as "Southeast" and distances such as "for four kilometers" are computed by examining the focus of the current proposition with respect to the previous spatial focus.

Finally, constraints on the realization of deictic anaphora (e.g., selecting "here" versus "there", "this" versus "that") can be related to SF, analogous to the use of DF to guide pronominalization. One view is to base deictic anaphora on the relation of the speaker's physical location to that of specificand (e.g., "I am here, you are there"). However, choice sometimes seems to be based on the relation of the specificand to the SF (e.g., substituting "here" in the second utterance of the above route plan). This is an issue of current investigation.

Limitations, Future Work

Despite the advantages of topical and temporal order in the LACE report, it remains simply a recounting of salient events. Even though it has a setting, it is not a story because it fails to explicitly indicate causal relations among events and states. Current research is focusing on analyzing the causal structure of short stories in an attempt to develop a theory of story structure/plot and its relationship to the event/state ontology (i.e., issues of motivation, enablement, causation, effect, and purpose, as well as spatiality and temporality).

Another area for research is the explicit representation of narrative techniques (e.g., suspense, surprise, mystery) as text plans. Because the text plans explicitly indicate what cognitive/psychological effect their application has on the hearer, they should be readily extensible to account for (simple) narrative techniques. Finally, more research needs to be done to investigate the various linguistic aspects of temporal and spatial adverbials and their relation to the event/state structure and text planning. In particular, temporal and spatial shift rules require more detailed examination.

Conclusion

This paper indicates how discourse focus, temporal focus, and a new notion of spatial focus can be used by a text planner to select and order content. In addition, the paper discusses how discourse, temporal, and spatial focus caches can be used to guide surface choices (e.g., pronominalization, the generation of tense and aspect, and the production of temporal and spatial adverbials). The use of these focal models in conjunction with narrative text plans allows the system to generate multiparagraph reports about the activities of a knowledge based simulator as well as locative instructions from a route planner.

Acknowledgments

I am grateful to Karen Sparck Jones and Steve Pulman of Cambridge University for many stimulating discussions as well as to the reviewers for useful feedback.

References

- [Allen, 1984] Allen, J. F. 1984. "Towards a General Theory of Time and Action." *Artificial Intelligence* 23(2):123-154.
- [Allen, 1988] Allen, J. F. editor, June, 1988. "Special Issue on Tense and Aspect." *Computational Linguistics* 14(2):
- [Anken, 1989] Anken, C. S. June, 1989. "LACE: Land Air Combat in Eric." Rome Air Development Center DL-9-0043.
- [Ehrich, 1987] Ehrich, V. 1987. "The Generation of Tense." *Natural Language Generation: Recent Advances in Artificial Intelligence, Psychology and Linguistics*, editor G. Kempen. 423-440. Dordrecht: Martinus Nijhoff.
- [Grosz, 1977] Grosz, B. J. "The Representation and Use of Focus in a System for Understanding Dialogs." *Proceedings of the Fifth Annual IJCAI*, Cambridge, MA, 1977. 67-76.
- [Grosz and Sidner, 1986] Grosz, B. J. and C. Sidner. July-September, 1986. "Attention, Intentions, and the Structure of Discourse." *Computational Linguistics* 12(3):175-204.

- [Hilton, 1987] Hilton, M. July, 1987. "ERIC: An Object-Oriented Simulation Language." Rome Air Development Center TR-87-103.
- [Hilton and Anken, 1990] Hilton, M. L. and C. S. Anken. February, 1990. "Map Display System: An Object-Oriented Design and Implementation." Rome Air Development Center Technical Report 90-54.
- [Hovy, 1988] Hovy, E. 1988.. "Planning Coherent Multisentential Text." Proceedings of the 26th Meeting of the ACL, Buffalo, NY, June 7-10, 1988. 163-169.
- [Hovy and McCoy, 1989] Hovy, E. and K. McCoy. 1989. "Focusing Your RST: A Step Toward Generating Coherent Multisentential Text." Submitted to Conference of Cognitive Science Society, March 20, 1989.
- [Kalita, 1989] Kalita, J. 1989. Automatically Generating Natural Language Reports. *International Journal of Man-Machine Studies* 30(1989):399-423.
- [Kittredge *et al.*, 1986] Kittredge, R., A. Polguère and E. Goldberg. 25-29 August, 1986. "Synthesizing Weather Forecasts from Formatted Data." Proceedings of COLING-86, The 11th International Conference on Computational Linguistics, University of Bonn, West Germany, 1986. 563-565.
- [Linde and Labov, 1975] Linde, C. and Labov, W. 1975. "Spatial Networks as a Site for the Study of Language and Thought." *Language*, 51(1975) 924-939.
- [Mann and Thompson, 1987] Mann, W. and S. Thompson. July, 1987. "Rhetorical Structure Theory: A Theory of Text Organization." Information Sciences Institute RS-87-190.
- [Matthiesen, 1984] Matthiesen, C. November, 1984. Choosing Tense in English, USC/ISI TR RR-84-143.
- [Maybury, 1989] Maybury, M. T. 1989b. "GENNY: A Knowledge-Based Text Generation System." *International Journal of Information Processing and Management* 25(2):137-150.
- [Maybury, 1990] Maybury, M. T. 1990. "Custom Explanations: Exploiting User Models to Plan Multisentential Text." Proceedings of the Second International Workshop on User Models, University of Honolulu, Hawaii, 30 March - 1 April, 1990.
- [McCoy and Cheng, 1988] McCoy, K. F. and J. Cheng. 1988. "Focus of Attention: Constraining what can be said next." Presented at the Fourth International Workshop on Text Generation, Catalina Island, Los Angeles, CA, 1988.
- [McKeown, 1982] McKeown, K. R. 1982. Generating Natural Language Text in Response to Questions About Database Structure. University of Pennsylvania PhD dissertation, TR MS-CIS-82-5.
- [Moens and Steedman, 1988] Moens, M. and M. Steedman. June, 1988. "Temporal Ontology and Temporal Reference." *Computational Linguistics* 14(2):15-28.
- [Moore, 1989] Moore, J. D. November, 1989. A Reactive Approach to Explanation in Expert and Advice-Giving Systems. Ph. D. dissertation, University of California at Los Angeles.
- [Nakhimovsky, 1988] Nakhimovsky, A. June, 1988. "Aspect, Aspectual Class, and the Temporal Structure of Narrative." *Computational Linguistics* 14(2):29-43.
- [Reichenbach, 1947] Reichenbach, H. 1947. *The Elements of Symbolic Logic*. London: Macmillan.
- [Searle, 1969] Searle, J. R. 1969. *Speech Acts*. Cambridge University Press.
- [Searle, 1975] Searle, J. R. 1975. "Indirect Speech Acts." *Syntax and Semantics 3: Speech Acts*, editors P. Cole and J. L. Morgan. 59-82. New York: Academic Press.
- [Sidner, 1979] Sidner, C. L. 1979. Toward a Computational Theory of Definite Anaphora Comprehension in English Discourse. Ph D dissertation, Massachusetts Institute of Technology, Cambridge, MA.
- [Sidner, 1983] Sidner, C. L. 1983. "Focusing in the Comprehension of Definite Anaphora." *Computational Models of Discourse*, editors M. Brady and R. Berwick. 267-330. Cambridge, MA: MIT Press.
- [Webber, 1987], B. L. 1987. "The Interpretation of Tense in Discourse." Proceedings of the 25th Annual Conference of the ACL, Stanford University, Palo Alto, CA, 1987. 147-154.
- [Winograd, 1972] Winograd, T. 1972. "Understanding Natural Language." from *Cognitive Psychology* 3(1). Orlando, Florida: Academic Press.