Discontinuities in Narratives

Alexander Nakhimovsky
Department of Computer Science
Colgate University
Hamilton, NY 13346

ABSTRACT

This paper is concerned with heuristics for segmenting narratives into units that form the basic elements of discourse representations and that constrain the application of focusing algorithms. The following classes of discontinuities are identified: figure-ground, space, time, perspective, and topic. It is suggested that rhetorical relations between narrative units are macro labels that stand for frequently occurring clusters of discontinuities. Heuristics for identifying discontinuities are presented and illustrated in an extended example.

1. The Segmentation Problem.
1.1. Introduction

This paper is concerned with heuristics for segmenting narratives into units that form the basic elements of discourse representations and that constrain the application of focusing algorithms. The importance of proper segmentation is frequently mentioned; as one text says, "the need for segmentation is almost universally agreed upon..." (Allen 1987: 398-399). However, "there is little consensus on what the segments of a particular discourse should be or how segmentation could be accomplished. One reason for this lack of consensus is that there is no precise definition of what a segment is beyond the intuition that certain sentences naturally group together" (ibid.). A brief discussion of our position is therefore in order, addressing the following questions: What is it that is being segmented? What is the dominant feature of a segment? What is the purpose of the segmentation process?

We adopt the position that reading a narrative, like taking part in a conversation, is a form of social interaction. However, the nature of the interaction is quite different in narratives and conversations, and so are the principles of segmentation and the nature of the resulting segments. The time of a conversation coincides with the time of its content. Simultaneity in time is typically accompanied by a tight integration between diachronic and synchronic behavior: the verbalization of how-to-get-there directions is the action of giving directions, task-oriented conversations between an expert and apprentice are an integral part of performing the task at hand, and the unfolding text of an argumentative dialogue is precisely the activity of arguing. Conversation can thus properly be called performative discourse. By contrast, the content of a narrative is decoupled from the linear progression of its text and unfolds in its own, separate timeline.

It follows that in place of the situation of discourse, a narrative is processed with respect to a constantly maintained deictic center, which is "the locus in conceptual space-time of the objects and events depicted or described by the sentences currently being perceived. At any point in the narrative, the cognitive agent's attention is focused on particular characters (and other objects) standing in particular temporal and spatial relations to each other. Moreover, the agent 'looks' at the narrative from the perspective of a particular character, spatial location and temporal location. Thus the deictic center consists of a WHERE-point, a WHEN-point and a WHO-point." (Bruder et al. 1986: 1). In this paper, the WHEN-point of the deictic center is referred to as the Temporal Focus (cf. Webber 1987a,b; Nakhimovsky 1987b).

We conceive of discourse segments (DSs) as continuous stretches of text corresponding to relatively monolithic pieces of internal representation. What "relatively monolithic" means is the subject of much of the rest of the paper; at this point, we simply say that the DS remains the same as long as the deictic center does not undergo drastic changes in space, time, perspective or composition, while the beginning of a new DS is accompanied/signaled by a discontinuity in one or several of these parameters. Within each segment, reference and anaphora are processed by local algorithms mostly relying on recency lists as in Sidner (1983). At the beginning of a new DS, a more global search through the accumulated representation is required.

We thus have three kinds of entities organized into three kinds of structures: linearly ordered stretches of text forming the Linear Text Structure (LTS); the Event-Situation Structure (ESS, cf. Webber 1987b), representing the narrative's unfolding contents; and the Current Focus Space, which is a collection of focusing mechanisms (including the deictic center) that together represent the "attentional state" (Grosz & Sidner 1986) of the system. The components of the LTS are frequently linked by rhetorical relations such as elaboration, resumption or flashback (see, e.g., Hobbs 1982). We believe that these rhetorical relations are simply macro labels that stand for certain oft-repeated clusters of discontinuities in the ESS. It is the discontinuities that are essential for constructing the ESS; the rhetorical labels need not be recognized by the reader at all, just as, on the sentence
level, speech acts need not be recognized in order to understand the intention of the speaker (Cohen & Levesque 1987).

The foregoing has served to motivate the need for reliable segmentation heuristics. Most such heuristics found in the literature are syntactical in nature, relying, in almost Eliza-like fashion, on clue words and phrases (see references in Grosz & Sidner 1986: 177). We propose that heuristics should be based on semantical considerations such as discontinuities in the representation. This paper investigates four kinds of discontinuities: discontinuities of topic, discontinuities of space and time, discontinuities of figure and ground, and discontinuities of the narrative perspective. Section 2 explains what these are; the remainder of this section gives a preliminary and partial illustration and reviews related work.

1.2 An example

Consider the following example:

(1) (a) Hartley and Phoebe had been sent by their mother to fix the tail vane of the windmill. (b) In the great expanse of the prairie where they lived, the high tower of the windmill was the only real landmark. (Worline 1956: 1)

Rhetorically speaking, sentence (b) interrupts a sequence of events described in (a) (= DS1) to start DS2, a description. In order to recognize this rhetorical relation between the two DSs, it is necessary to recognize that:

(1) there is a shift of topic;
(2) there is a shift in perceptual modality to visual perception;
(3) there is a shift in time scale, from the events of the current day to years or decades, associated with the lifetime of a windmill and the 'where they lived' clause;
(4) there is shift in spatial scale from a household to the entire prairie;
(5) there is a shift from a foregrounded sequence of events to the 'mopping-up' operation of filling in the background; the shift is signaled by aspectual changes (Hopper 1978; Nakhimovsky 1987b, 1988); and
(6) the foregrounded sequence of events includes two telic processes (walking from the house to the windmill; fixing the broken part) whose beginnings or preconditions are explicitly mentioned but whose terminal points are still in the future relative to the Temporal Focus; this sets up expectations for the reader.

The following empirical investigation is suggested by this and similar examples: classify the discontinuities and clusters of discontinuities that typically accompany DS breaks in narratives; identify the linguistic and extra-linguistic knowledge involved; develop heuristics for using this knowledge; and test the heuristics in a computer program. Section 2 below discusses narrative representations and the data structures that are needed for constructing and updating them. A classification of discontinuities falls out of this discussion. Section 3 presents several kinds of knowledge that we consider relevant for identifying discontinuities in narratives. Section 4 illustrates the corresponding heuristics with an extended example whose point is that our heuristics explain the use of deictic devices and correctly predict paragraph breaks observed in existing narratives.1

1.3. Related work

Central to our investigation is the idea that text understanding involves "building a good structure" (Bruce 1981: 283) by the process of successive embedding of the current sentence representation into the structure created by the preceding text. This idea emerged almost simultaneously in Computational Linguistics/AI (Bruce 1981; Webber 1979, 1983) and linguistics (Kamp 1981, Heim 1982). (There has also been interaction, more or less conscious, with the Reader-Response school of literary criticism as represented in, e.g., Thompson 1981.) The main difference between these two developments concerns the role of inference and monotonicity: to what extent does the structure being built incorporate defeasible inferences that may have to be undone? We do not pursue this issue here but assume non-monotonic embeddability and an active role for expectations set up by the text. (Cf. Nakhimovsky 1988 for some discussion.)

Grosz & Sidner (1986) is the first unified approach in which the problem of segmentation is tied up with the notion of attentional state. As argued in Nakhimovsky (1987b), some features of their model (a stack mechanism for attentional state, the prominence of pragmatic notions such as the speaker's intentions) make it more appropriate for conversation rather than narrative. Webber (1988) and Nakhimovsky (1988) suggest a model in which the distinction between the top and the rest of the stack is replaced by the distinction between a short-term memory and the ESS.

Our work has been developing in close contact with the SUNY Buffalo Graduate Group in Cognitive Science's project on cognitive and computer systems for understanding narrative text. This research program consists of a group of projects whose goals are to develop a psychologically real model of a cognitive agent's comprehension of deictic information in narrative text. The hypothesis of this project is that the construction and modification of the deictic center is important for comprehension. To test this hypothesis, a computer system implemented in SNePS (Shapiro 1979, Shapiro & Rapaport 1987) is being developed that will "read" a narrative and answer questions concerning the reader's beliefs about the objects, relations, and events in it. The system will be psychologically real, because the details of the algorithms and the efficacy of the linguistic devices will be validated by psychological experiments on normal and abnormal comprehenders. (Cf. Bruder et al. 1986, Daniels 1986, Almeida 1987, Wiebe & Rapaport 1988.)

There is surprisingly little psychological work on discourse segmentation. To remedy this situation, a series of experiments

1 This is not to suggest that typographical paragraph breaks are the only discontinuities we are after. First, at the minimal level of segmentation, DSs are usually smaller than typographical paragraphs. Second, a typographical paragraph does not simply signal or suggest a discontinuity: it creates one by its very presence.
is being designed and conducted at Colgate (Reynolds & Nakhi-

mowsky (in preparation)). The subjects are being asked to

segment narratives of various genres and structural characteris-
tics and later recall them, under conditions of both cued and

uncued recall. We proceed from the hypothesis that narrative

structures as they are remembered are different from such struc-
tures as they are built "on line" in the process of comprehension.

2. Narratives and Their Representations.

For the purposes of this paper, we adopt the simplest narra-
tive model. There is only one, objective, narrator, who gives an

observer account of the story. The story consists of descriptions

of situations evolving or persisting in time. Loosely following

Hayes (1983), we call these descriptions history-tokens, or h-
tokens for short. (Lexical meanings are history-types (h-types)

related to their h-tokens by the relation of "non-monotonic

embeddability".) The time of a narrated h-token is determined

with respect to a specific point within the previously narrated h-
tokens, and usually without reference to the time of discourse.‡

Using Reichenbach's notion of reference time, we can say that

a sentence in a narrative never has the time of discourse as its

reference time. Put differently, a sentence in a present tense

necessarily attempts a narrative, unless it is interpreted as histor-

ical (narrative) present.

The meaning of a non-generic sentence in a narrative is a

collection of h-tokens together with the position of the deictic

center (to the extent that it is uniquely specified by the sentence).

After the first sentence of a narrative establishes the initial set-up,

each subsequent sentence is processed in the context consisting

of the ESS and the Current Focus Space. Depending on the

meaning of the sentence, one of two things happens: either the

representation of the sentence is incorporated in the Current

Focus Space, with the focusing mechanisms appropriately

modified, or, in the case of a focus shift, the contents of the

Current Focus Space are incorporated into the Event-Situation

Structure (ESS) and Linear Text Structure (LTS), and the Current

Focus Space is completely react. The nature of the processing at

the DS juncture is thus quite different from the "routine" tasks to

be performed as long as the text remains in the same DS: the

start of a new DS prompts, and is prompted by, a shift of atten-

tion. The circularity here is deliberate. The start of a new DS

brings about several changes, some of them more immediately

noticeable than others; the more obvious ones serve to signal that

a new DS is, indeed, started. We proceed to catalog the changes.

The first observation to make is that a narrative must have a

plot, i.e., present a sequence of events that forms an instance of a

recognizable pattern. (The patterns are partly universal, partly

culture-specific; the work of Lehneri (1982) and Alker et al.

(1985) can be seen as a search for the principles on which such

patterns are built.) Using Gestalt terminology (brought into

linguistics by Talmy (1983)), we can say that a narrative's plot

must present a recognizable temporal/cause figure shown against

some ground that minimally consists of spatial/visual settings

(descriptions of characters are also frequent). The distinction is

not always clear-cut, because elements of the figure can be hid-

den among the details of the ground, but the temporal nature of

the plot does stand in clear contrast to the spatial nature of the

ground.

Secondly, a narrative must have characters with whom we

empathize. These characters don't have to be human: one can

easily imagine a story about an adventurous seed that falls off its

parent, gets swallowed and excreted by a horse, and

nearly drowns in a tropical rain before being miraculously saved

by the sun and producing a flower. Even so, the narrative is

likely to alternate between the objective narrator's point of view

and that of one of the characters. ("The belly of the horse was
dark and noisy inside.")

Suppose for a moment that the ESS is implemented as a par-
titioned network of nodes, each node representing a narrated h-
token and partitions corresponding to the belief spaces of the

characters. An unfolding narrative then creates a path through

the network such that it, and the inferences it generates, covers

the network in its entirety. Some discontinuities of narratives

evolve simply reflect the intrinsic spatial or temporal discontinui-

ties of the plot, e.g., when the story is composed of a sequence

of events taking place during an afternoon, followed by two years' 
hiatus, followed by another action-packed afternoon. (A similar

element of a spatial discontinuity can be easily imagined.) One

measure of the "simplicity" of a narrative is how faithfully the

order and structuring of its text reflects the order and structuring

of its component events. Even in the simplest narrative, how-

ever, there are bound to be discontinuities resulting from the ten-

tion between the linear nature of the text and the multi-
dimensional structure that it is meant to evoke. These are discon-

tinuities of figure and ground, when the narrative shifts between

the main story-line and the surrounding circumstances, and

discontinuities of perspective, when the narrative crosses into a
different "empathy partition" or creates a new one.

Given this classification of discontinuities, one can proceed

to catalog the clues that signal them. This is a subject for a large

empirical study, of which the next section is but a preliminary

sketch. It is important to keep in mind that it is clusters of
discontinuities that signal the beginning of a new DS.

3. Discontinuities in Narratives.

3.1. Topic discontinuities.

Discontinuities of topic fall into two groups. In the first,

there is no anaphoric relation or immediate inference path from

the new topic to a node in the Current Focus Space. What an

"immediate inference path" is depends, of course, on the

system's knowledge base and inferential capabilities, but this is a
related phenomena such as a change in temperature or lighting:

(2) [The barmbracks] had been cut into long thick slices and were ready to be handed round at tea. Maria had cut them herself.

Maria was a very, very small person indeed . . .

3.2. Temporal discontinuities.

The most important temporal discontinuities are:

a. A shift from perfective to imperfective sentence perspective accompanied by a shift to a much greater time scale. The corresponding rhetorical move is frequently characterized as "introducing background or descriptive material". The move is frequently accompanied by a topic re-introduction.

b. The reverse shift from descriptive material to the main line of the narrative. This move is signaled by the TF and the entire deictic center, returning to an established node in the ESS, with an appropriate contraction of the time scale.

c. A backwards move of the TF to an earlier point in time, with or without a change in time scale. Rhetorically, this is known as flashback. This move is frequently signaled by a verb in past perfect or by the used to + Infinitive construction, although a shift may occur without such a verb form, and the presence of such a verb form does not necessarily signal a shift: the reference time of the sentence may remain the same as, rather than precede, the current TF. (See Almeida (1987) and Nakhimovsky (1988) for a more detailed discussion.)

3.3. Spatial discontinuities.

The most obvious spatial discontinuities are discontinuities of scale. It is argued at some length in Nakhimovsky (1986, 1987a) that h-types have time scales associated with them, characterized in terms of "received" cyclical events such as day or year. It seems equally necessary to establish a gradation of spatial scales, based on similar considerations from human biology and habitat. The spatial scales we currently employ are: one's body (which may need to be further specialized); within arm's reach; room area, such as desk or bed; room; floor; house; household; village/neighborhood; larger area within a between-meals round trip; within a day's round trip; staying overnight. (The larger the scale, the more domain- and culture-specific variation there is.)

A change in spatial scale is frequently accompanied by related phenomena such as a change in temperature or lighting:

(3) She changed her blouse too and, as she stood before the mirror, she thought of how she used to dress for mass on Sunday morning when she was a young girl; and she looked with quaint affection at the diminutive body which she had so often adorned. In spite of its years she found it a nice tidy little body.

When she got outside the streets were shining with rain and she was glad of her old brown raincloak.

3.4. Perspective discontinuities.

The following signals frequently indicate a shift to a character's subjective perspective away from the objective perspective of the implied narrator: (a) attitude reports, which can be further subcategorized into beliefs, desires, emotions, and so on; communication verbs; and perceptions; and (b) deictics. Attitude reports, by their very nature, introduce private semantic objects into the ESS and thus create, or move into, a partition. Deictics, such as the verbs bring or come, indicate the position of the WHO-point of the deictic center; a shift of this position to one of the characters signals a discontinuity. (Cf. Banfield 1982, Wiebe & Rapaport 1988.)

4. An Example.

This section of the paper applies the above heuristics to the initial fragment of Joyce's "Clay". The fragment falls into three DSs: sentence 1 (DS1), sentences 2-6 (DS2), and sentences 6-12 (DS3). Embedded in DS3 is DS3.1, consisting of sentences 9-10. Sentence 13 starts a fourth DS by returning to the first one.

1. The matron had given her leave to go out as soon as the women's tea was over and Maria looked forward to her evening out. The kitchen was spick and span: the cook said you could see yourself in the big copper boilers. The fire was nice and bright and on one of the side-tables were four very big barmbracks. These barmbracks seemed uncut; but if you went closer you would see that they had been cut into long thick even slices and were ready to be handed round at tea. Maria had cut them herself.

6. Maria was a very, very small person indeed but she had a very long nose and a very long chin. She talked a little through her nose, always soothingly: Yes, my dear, and No, my dear. She was always sent for when the women quarrelled over their tubs and always succeeded in making peace. One day the matron had said to her:-Maria, you are a veritable peace-maker! And the sub-matron and two of the Board ladies had heard the compliment. And Ginger Mooney was always saying what she wouldn't do to the dummy who had charge of the irons if it wasn't for Maria. Everyone was so fond of Maria.

13. The women would have their tea at six o'clock and she would be able to get away before seven.

The transition from sentence 1 to 2 is very similar to the transition from (a) to (b) in example (1): it is a shift from figure to ground marked primarily by tense and aspect changes and a shift from temporal to spatial/visual material. The change in time and space scales is not as dramatic here, but there is the

†These intuitive divisions, two of which correspond to Joyce's paragraph breaks, have been confirmed by one of the experiments reported in Reynolds and Nakhimovsky (in preparation).
same arrangement of narrated events that are prior to the Temporal Focus ("the matron had given her leave...") and anticipated events that are in the future with respect to the Temporal Focus (the women's tea, Maria's evening out). This sets up expectations suspended by the visual material and stative verbs.

Sentence 5, by using an action verb and the past-perfect tense, return to the time scale and the temporal arrangement of sentence 1.

The transition from 5 to 6 (paragraph break) is characterized by a shift in time scale and a topic re-introduction. The material is a (back)ground character description: 'Maria is such that...'. The beginning of DS3.1 is signaled by a well-known clue phrase *one day* and by the past-perfect tense. Note that the material is still ground ('Maria is such that...'). The precise relationship between DS3 and DS3.1 is at this point ambiguous: they can be siblings, and they would be if sentences 11 and 12 were dropped from the narrative. However, sentence 11 clearly signals a return to the material of sentences 6-8: the tense changes back from past perfect to past (past progressive, presumably for a sharper contrast with the perfect), and the adverb *always*, used in sentences 7 and 8, reappears again. This establishes that DS3.1 is, indeed, embedded in DS3; note that it bears no relation to DS1, and, in particular, the past-perfect events of sentences 9-10 are much further in the past (on a different time scale) than the past perfect events of sentence 1.

The transition from 12 to 13 (the next paragraph break) is a return to the event sequence of sentence 1. The position of the TF, and the entire deictic center is indicated by the future-in-the-past tense. Note that it is essential to retrieve the entire deictic center, and not just the TF, because the WHO-point is also restored: the *she* in 13 does not evoke *Maria* of sentence 12 or any other sentence in DS 2 and 3 presented from the perspective of the implied narrator. Rather, this pronoun is a quasi-indexical (Rapaport 1986) that replaces the first-person singular *I* of Maria's thoughts and expectations, signaled by *Maria looked forward* in sentence 1.†

5. Future research.

There are several directions in which we are proceeding. Wiebe & Rapaport (1988) and Wiebe (in progress) present an outline of a detailed computational investigation of narrative perspective and reference. Reynolds & Nakhimovsky (in preparation) will report on several psychological experiments designed to obtain empirical data on how people segment narratives in the process of reading, and how they are segmented when recalled. The deictic center project (Bruder et al. (1986)) contains both a linguistic study of the role of indexicals in narrative segmentation and a computational project that will test all the diverse segmentation heuristics within the uniform system of belief representation.

Acknowledgments.

This research was supported in part by NSF Grant No. IRI-8610517. We are grateful to Mary Galbraith and Janice M. Wiebe for comments on an earlier version, and to Bonnie Webber and Rebecca Passonneau for discussions of related work.

References.

Alber, E., et al. 1985. Two reinterpretations of Toynbee's Jesus: Explorations in computational hermeneutics. In G. Tonfoni (ed.), Artificial Intelligence and Text Understanding. Quaderni di Ricerca Linguistica 6: 49-94.

Allen, James. 1984. Towards a general theory of action and time, Artificial Intelligence 23: 123-154.

Allen, James. 1987. Natural Language Understanding. Menlo Park, CA: Benjamin Cummings Publishing Co.

Aimeida, Michael J. 1987. Reasoning about the temporal structure of narratives. Technical Report 87-10. Buffalo: SUNY Buffalo Department of Computer Science.

Banfield, Ann. 1982. Unspaukable Sentences: Narration and Representation in the Language of Fiction. Boston: Routledge & Kegan Paul.

Bruce, Bertram. 1981. A social interaction model of reading. Discourse Processes 4: 273-311.

Bruder, G. A., et al. 1986. Deictic centers in narrative: An interdisciplinary cognitive-science project. Technical Report 86-20. Buffalo, NY: SUNY Buffalo Department of Computer Science.

Cohen, Philip R., & Levesque, Hector J. 1987. Persistence, intention and commitment. CSLI-87-88. Stanford, CA: CSLI.

Daniels, Joyce H. 1986. A psychological investigation into the deictic center. Proc. 8th Annual Conference of the Cognitive Science Soc. (Univ. of Mass. at Amherst). Hillsdale, NJ: Lawrence Erlbaum: 621-626.

Grosz, Barbara, & Sidner, Candace. 1986. Attention, intentions, and the structure of discourse. Computational Linguistics 12: 175-204.

Hayes, P. 1985. The second naive physics manifesto. In J. Hobbs & R. Moore (eds.), Formal Theories of the Commonsense World. Norwood, NJ: Ablex: 1-36.

Heim, I. 1982. The semantics of definite and indefinite noun phrases. Ph.D. Dissertation, Department of Linguistics, University of Massachusetts, Amherst.

Hinrichs, Erhard. 1986. Temporal anaphora in discourses of English. Linguistics and Philosophy, 9: 63-82.
