



University of Pennsylvania
ScholarlyCommons

Technical Reports (CIS)

Department of Computer & Information Science

4-1-1988

Discourse Deixis: Reference to Discourse Segments

Bonnie L. Webber

University of Pennsylvania, bonnie@inf.ed.ac.uk

Follow this and additional works at: https://repository.upenn.edu/cis_reports

Recommended Citation

Bonnie L. Webber, "Discourse Deixis: Reference to Discourse Segments", . April 1988.

University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-88-27.

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/cis_reports/458
For more information, please contact repository@pobox.upenn.edu.

Discourse Deixis: Reference to Discourse Segments

Abstract

Computational approaches to discourse understanding have a two-part goal: (1) to identify those aspects of discourse understanding that require process-based accounts, and (2) to characterize the processes and data structures they involve. To date, in the area of reference, process-based accounts have been developed for subsequent reference via anaphoric pronouns and reference via definite descriptors. In this paper, I propose and argue for a process-based account of subsequent reference via deictic expressions. A significant feature of this account is that it attributes distinct mental reality to units of text often called *discourse segments*, a reality that is distinct from that of the entities described therein.

Comments

University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-88-27.

**DISCOURSE DEIXIS:
REFERENCE TO DISCOURSE
SEGMENTS**

Bonnie Lynn Webber

**MS-CIS-88-27
LINC LAB 110**

**Department of Computer and Information Science
School of Engineering and Applied Science
University of Pennsylvania
Philadelphia, PA 19104**

April 1988

**To appear in Proceedings of the 26th Annual Meeting of the Association
for Computational Linguistics, June 1988, SUNY at Buffalo.**

Acknowledgements: This research was supported in part by ARO grants DAA29-84-K-0061, DAA29-84-9-0027, NSF grant MCS-8219196-CER, DARPA grants N00014-85-K-0018, N00014-85-C-0012, and an Alvey grant to the Centre for Speech Technology Research, University of Edinburgh. It was completed while visiting at Medical Computer Science, Stanford University.

DISCOURSE DEIXIS: REFERENCE TO DISCOURSE SEGMENTS

Bonnie Lynn Webber
Department of Computer & Information Science
University of Pennsylvania
Philadelphia PA 19104-6389

ABSTRACT

Computational approaches to discourse understanding have a two-part goal: (1) to identify those aspects of discourse understanding that require process-based accounts, and (2) to characterize the processes and data structures they involve. To date, in the area of reference, process-based accounts have been developed for subsequent reference via anaphoric pronouns and reference via definite descriptors. In this paper, I propose and argue for a process-based account of subsequent reference via deictic expressions. A significant feature of this account is that it attributes distinct mental reality to units of text often called **discourse segments**, a reality that is distinct from that of the entities described therein.

1. INTRODUCTION

There seem to be at least two constructs that most current theories of discourse understanding have adopted in at least some form. The first is the **discourse entity**, first introduced by Lauri Karttunen in 1976 (under the name "discourse referent") [9] and employed (under various other names) by many researchers, including myself [18]. The other is the **discourse segment**.

Discourse entities provide these theories with a uniform way of explaining what it is that noun phrases (NP) and pronouns in a discourse refer to. Some NPs **evoke** a new discourse entity in the listener's evolving model of the discourse (which I have called simply a **discourse model**), others refer to ones that are already there. Such entities may correspond to something in the outside world, but they do not have to. To avoid confusion with a sense of "referring in the outside world", I will use the

terms **refer_m** here, meaning "refer in a model", and **referent_m**, for the entity in the model picked out by the linguistic expression.

The basic features of a discourse entity are that (a) it is a constant within the current discourse model and that (b) one can attribute to it, *inter alia*, properties and relationships with other entities. (It is for this reason that Bill Woods once called them "conceptual coat hooks".) In some theories, different parts of the discourse model (often called **spaces**) may represent different modalities, including hypothetical contexts, quantified contexts, the belief contexts of different agents, etc. Depending on what space is currently being described, the same NP or pronoun may evoke and/or **refer_m** to very different discourse entities.

The other common construct is the **discourse segment**. While discourse segmentation is generally taken to be a chunking of a linguistic text into sequences of related clauses or sentences, James Allen notes:

... 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 [[1], p. 398-9]

What is taken to unify a segment is different in different theories: for example, among computational linguists, Grosz & Sidner [5] take a discourse segment to be a chunk of text that expresses a common purpose (what they have called a **discourse segment purpose**) with respect to the speaker's plans; Hobbs [8] takes a discourse segment to be a chunk of text that has a common meaning; while Nakhimovsky [12], considering only narrative, takes a discourse segment to be a chunk of text that describes a single event from a single perspective.

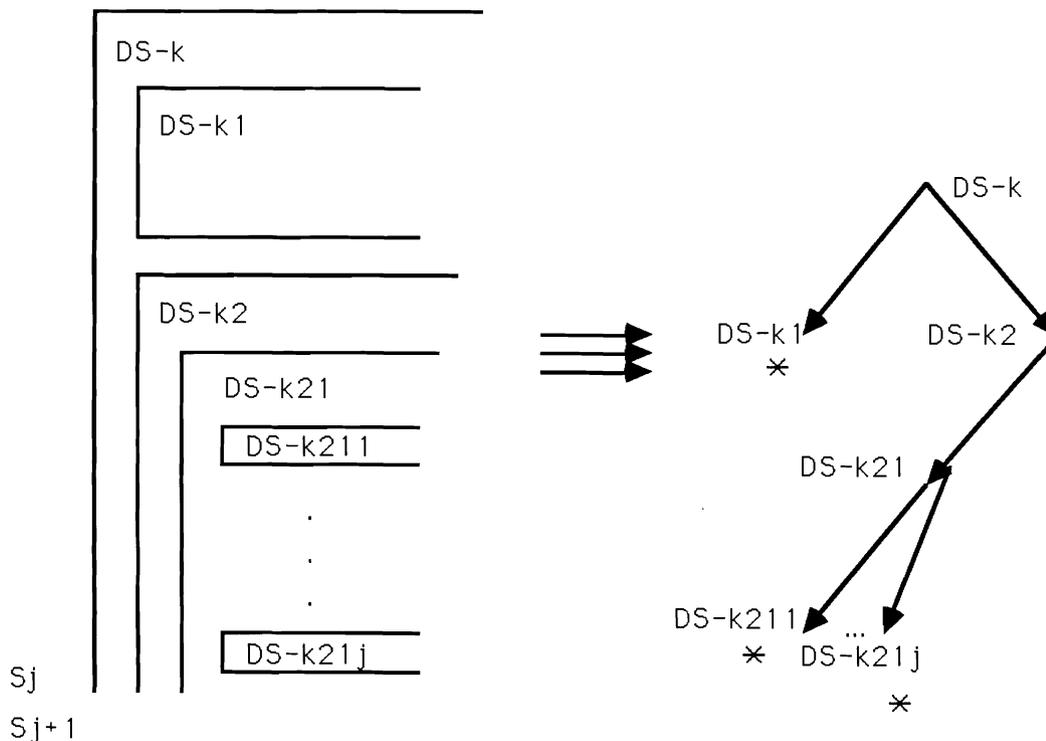


Figure 1. Discourse Segmentation

While **discourse segment** is usually defined recursively, theories differ in what they take the minimal segment to be. Hobbs takes it to be a **sentence**, and Polanyi [12], a **clause**. Grosz & Sidner do not state explicitly how much is needed to express a single purpose, but from their examples, it appears to be a single sentence as well. (Unlike Hobbs and Polanyi, Grosz & Sidner do not consider every sentence to be a discourse segment *per se*.) Since discourse segment is defined recursively, the resulting segmentation of a text (or at least, large parts of it) can be described as a tree. From the point of view of processing, this means that at any point in the discourse, several segments, each embedded in the one higher, may still be **open** - i.e., under construction. This is illustrated schematically in Figure 1.

At the point of processing sentence S_{j+1} in this example, segments DS_{k1} , DS_{k211} , ..., DS_{k21j} are complete (**closed** - indicated by a *), while DS_k , DS_{k2} , and DS_{k21} are **open**, able to incorporate sentence S_{j+1} (or, alternatively, its corresponding unary discourse segment). Of special interest is the **right frontier** of the tree -- the set of nodes comprising the most recent closed segment and all currently open segments -- here $\{DS_{k21j}, DS_{k21}, DS_{k2}, \text{ and } DS_k\}$, which I will make use of later in Section 3. Several researchers (including Grosz & Sidner [5], Hirschberg & Litman [6], Robin Cohen

[2] and Rachel Reichman [15]) have discussed problems inherent in this discourse parsing task, among which is the lack of precise definition of its basic building block.

For the current discussion, the most significant thing about these two constructs is their different associations: **discourse entities** go with NPs (to explain anaphoric and definite reference_m) and **discourse segments** go with sentences or clauses (to explain textual coherence and discourse structure). This leaves a gap in the case of reference_m to what can only be taken to be some aspect of a sequence of clauses, sentences or utterances (e.g., its content, form, modality, etc.), for example:

Example 1

It's always been presumed that when the glaciers receded, the area got very hot. The Folsom men couldn't adapt, and they died out. **That's** what is supposed to have happened. It's the textbook dogma. But it's wrong. They were human and smart. They adapted their weapons and culture, and they survived.

Example 2

The tools come from the development of new types of computing devices. Just as we thought of intelligence in terms of

servomechanism in the 1950s, and in terms of sequential computers in the sixties and seventies, we are now beginning to think in terms of parallel computers, in which tens of thousands of processors work together. **This** is not a deep, philosophical shift, but it is of great practical importance, since it is now possible to study large emergent systems experimentally. [[6] p.176]

The obvious question is whether such reference_m involves the same processes used to explain how a pronoun or NP evokes and/or refers_m to a discourse entity or whether some other sort of process is involved. In this paper I will argue for the latter, giving evidence for a separate reference_m process by which a linguistic expression is first interpreted as a pointer to the representation of a **discourse segment** and then further constrained to specify either (a) a particular aspect of the discourse segment (e.g., its form, interpretation, speech act, etc.) or (b) a particular entity within its interpretation.

In Section 2, I will attempt to justify the existence of a second referring_m process linked to a representation of discourse segments *per se*. In Section 3, I will attempt to justify particular features of the proposed process, and Section 4 summarizes the implications of this work for discourse understanding.

2. Justifying a Second Referring_m Process

There is ample evidence that subsequent reference can be made to some aspect of a sequence of clauses in text. Besides Examples 1 and 2 above, several other examples will be presented later, and the reader should have no trouble finding more. So the existence of such a phenomenon is not in dispute. Also not in dispute is the fact that such subsequent reference is most often done via deictic pronouns: Of 79 instances of pronominal reference_m to clausal material found in five written texts¹, only 14 (~18%) used the pronoun **it** while the other 65 (~82%) used either **this** or **that** (17 instances of **that** and 48 of **this**). On the other hand, looking at all instances of pronominal reference_m using **it** to discourse entities evoked by NPs², of 41 such references, 39 (~95%) used **it** while only 2 (~5%) used **this** or **that**. Because of this, I will call this type of reference **discourse deixis**.

The first thing to note about discourse deixis is that the referent_m is often distinct from the things described in the sequence. For example,

Example 3

There's two houses you might be interested in:

House A is in Palo Alto. It's got 3 bedrooms and 2 baths, and was built in 1950. It's on a quarter acre, with a lovely garden, and the owner is asking \$425K. But **that's** all I know about it.

House B is in Portola Vally. It's got 3 bedrooms, 4 baths and a kidney-shaped pool, and was also built in 1950. It's on 4 acres of steep wooded slope, with a view of the mountains. The owner is asking \$600K. I heard all **this** from a friend, who saw the house yesterday.

Is **that** enough information for you to decide which to look at?

In this passage, **that** in the second paragraph does not refer to House A (although all instances of it do): rather it refers to the description of House A presented there. Similarly (all) **this** in the third paragraph does not refer to House B (although again, all instances of it do): rather it refers to the description of House B presented there. **That** in the fourth paragraph refers to the descriptions of the two houses taken together.

That in each case it is the given description(s) that **this** and **that** are accessing and not the houses, can be seen by interleaving the two descriptions, a technique often used when comparing two items:

Example 4

There's two houses you might be interested in:

House A is in Palo Alto, House B in Portola Vally. Both were built in 1950, and both have 3 bedrooms. House A has 2 baths, and B, 4. House B also has a kidney-shaped pool. House A is on a quarter acre, with a lovely garden, while House B is on 4 acres of steep wooded slope, with a view of the mountains. The owner of House A is asking \$425K. The owner of House B is asking \$600K. **#That's** all I know about House A. **#This** I heard from a friend, who saw House B before it came on the market.

Is **that** enough information for you to decide which to look at?

Here houses A and B are described together, and the failure of **that** and **this** to refer successfully in the second paragraph indicates that (a) it is not the houses being referred_m to and (b) the individual descriptions available for reference_m in Example 3 are no longer available here. One must conclude from this that it is

something associated with the sequences themselves rather than the discourse entities described therein that **this** and **that** refer_m to here.

The next thing to note is that the only sequences of utterances that appear to allow such pronominal reference_m are ones that intuitively constitute a **discourse segment** (cf. Section 1), as in Example 1 (repeated here) and Example 5:

Example 1

It's always been presumed that [1 when the glaciers receded, the area got very hot. The Folsom men couldn't adapt, and they died out.1] **That's** what is supposed to have happened. It's the textbook dogma. But it's wrong. They were human and smart. They adapted their weapons and culture, and they survived.

Example 5

...it should be possible to identify certain functions as being unnecessary for thought by studying patients whose cognitive abilities are unaffected by locally confined damage to the brain. For example, [1 binocular stereo fusion is known to take place in a specific area of the cortex near the back of the head. [2 Patients with damage to this area of the cortex have visual handicaps but show no obvious impairment in their ability to think.2] **This** suggests that stereo fusion is not necessary for thought.1] **This** is a simple example, and the conclusion is not surprising.... [[6], p. 185]

In Example 1, **that** can be taken to refer_m to the narrative of the glaciers and the Folsom men, which is intuitively a coherent discourse segment. (Brackets have been added to indicate discourse segments. Subscripts allow for embedded segments.) In Example 5, the first **this** can be taken as referring to the observation about visual cortex-damaged patients. The second **this** can be taken as referring to the whole embedded "brain damage" example.

To summarize the current claim: in the process of discourse understanding, a referent_m must be associated with each discourse segment, independent of the things it describes. Moreover, as Example 6 shows, this referent_m must have at least three properties associated with it: the speech act import of the segment, the form of the segment, and its interpretation (e.g., as a situation, event, object description, etc.)

Example 6

A: Hey, they've promoted Fred to second vice president.

(* **that speech act** *)

B1: **That's** a lie.

(* **that expression** *)

B2: **That's** a funny way to describe the situation.

(* **that event** *)

B3: When did **that** happen?

(* **that action** *)

B4: **That's** a weird thing for them to do.

I have not said anything about whether or not these discourse segment referents_m should be considered discourse entities like their NP-evoked counterparts. This is because I do not believe there is enough evidence to warrant taking a stand. Part of the problem is that there is no precise criterion for "discourse entity-hood".³ However, if every discourse segment evokes a discourse entity, an account will be needed of (1) when in the course of processing a segment such a thing happens, and (2) what the 'focus' status of each of these entities is.

3. Features of Deictic Reference_m

I suggest that the process of resolving discourse segment reference_m involves the following steps:

1. An input pronoun is first interpreted as a pointer to a representation of a discourse segment on the right frontier (cf. Section 1).
2. As the rest of the clause containing the pronoun is interpreted, pronoun interpretation may be either
 - a. further constrained to some property of the discourse segment representation
 - b. extended to one of the discourse entities within the interpretation of the segment
3. As a consequence of whether **this** or **that** was used, the listener characterizes the speaker's "psychological distance" to its referent_m as either "close" or "far away". That is, this well-known deictic feature of **this/that** is not used in the referent-finding process but rather afterwards, in attributing the speaker's relationship to that referent_m.

In this section, I will try to motivate each of the proposed steps.

I have already argued that some deictic pronouns must be interpreted with respect to a discourse segment. Here I claim that the only discourse segments so available are ones on the right frontier. My evidence for this consists of (a) it being true of the 69 clausally-referring instances of **this** and **that** found in the five texts and (b) the oddity of examples like the following variation of Example 3 where **that** in paragraph 3 is intended to refer_m to the description of House A.

Example 3'

There's two houses you might be interested in:

House A is in Palo Alto. It's got 3 bedrooms and 2 baths, and was built in 1950. It's on a quarter acre, with a lovely garden, and the owner is asking \$425K.

House B is in Portola Vally. It's got 3 bedrooms, 4 baths and a kidney-shaped pool, and was also built in 1950. It's on 4 acres of steep wooded slope, with a view of the mountains. The owner is asking \$600K. I heard all **this** from a friend, who saw the house yesterday. #But **that's** all I know about House A.⁴

Is **that** enough information for you to decide which to look at?

(Note that this very limited availability of possible referents_m and the ability to coerce referents to any of their parts which I shall argue for shortly suggests parallels between this phenomenon and definite NP and temporal anaphora.)

Because at any time, there may be more than one discourse segment on the right frontier, part of the reference resolution process involves identifying which one is intended. To see this, re-consider the first part of Example 5.

Example 5

...it should be possible to identify certain functions as being unnecessary for thought by studying patients whose cognitive abilities are unaffected by locally confined damage to the brain. For example, binocular stereo fusion is known to take place in a specific area of the cortex near the back of the head. Patients with damage to this area of the cortex have visual handicaps but show no obvious impairment in their ability to think. **This**

At this point in the discourse, there are several things that **this** can be taken as specifying.

Considering just the things associated with clauses (and just this segment of text, and not what it is embedded in), **this** can be taken as specifying either the segment associated with the previous sentence (as in the original text - "**This** suggests that stereo fusion is not necessary for thought.") or the segment associated with the description of the whole example - "**This** is only a simple example, and the conclusion is not surprising..."). The listener's choice depends on what is compatible with the meaning of the rest of the sentence.⁵ As with other types of ambiguity, there may be a default (i.e. context-independent) preference for one particular form of construal over the others (cf. [3]) but it is easily over-ridden by context.

This ambiguity as to the intended designatum of a pointer is very similar to the ambiguity associated with the more fundamental and historically prior use of deixis in pointing within a shared spatio-temporal context, as in the following example:

Example 7

[A and A Junior are standing in A's art gallery]
A: Someday **this** will all be yours.

Here **this** could be interpreted as either the business, the pictures, or the physical gallery.⁶ Both Quine [14] and Miller [10] have observed in this regard that all pointing is ambiguous: the intended **demonstratum** of a pointing gesture can be any of the infinite number of points "intersected" by the gesture or any of the structures encompassing those points. (Or, one might add, any interpretation of those structures.) The ambiguity here as to how large a segment on the right frontier is encompassed by a **this** or **that** is very similar.

(Another feature that Quine and Miller mention, that will come up later in this discussion, involves constraints on the demonstratum of a pointing gesture to being something present in the shared context or some mutually recognizable re-interpretation of it. The latter is what Quine has called **deferred ostension**. It enables one, given the right audience, to point to the ceiling, with wires dangling from the center, say "That's off being cleaned" and effectively refer to the chandelier. Most examples of deferred ostension, both in spatio-temporal deixis and discourse deixis, are not that extreme. However, as I will try to show, both these features -- ambiguity and "required presence" -- are characteristic of discourse deixis as well.)

Having taken the initial step of interpreting a pronoun as pointing to the representation of a discourse segment, the proposed process must then be

able to further coerce [8,11] that interpretation to be some **property** of the discourse segment representation or to some **entity** within it. Example 6 (above) illustrates the first type of coercion, Example 8, the latter.

Example 8

A: In the Antarctic autumn, Emperor penguins migrate to Tasmania.

B₁: **That's** where they wait out the long Antarctic winter.

(* that place *)

B₂: So **that's** what you're likely to see there in May.

(* that species of birds *)

B₃: **That's** when it begins to get too cold even for a penguin.

(* that time *)

The reason for taking discourse segment identification and coercion as two separate steps in the process is to accommodate the fact that most instances of **this** and **that** are as the first NP in a clause.⁷ Since the listener cannot say for sure what they refer_m to until more evidence comes in from the rest of the sentence, a two-stage process allows the first stage of the process to be done immediately, with the second stage done as a subsequent constraint satisfaction process. This would resemble spatio-temporal uses of **this** and **that**, where the listener recognizes the general pointing gesture, and then tries to figure out the intended demonstratum based on what the speaker says about it (and on general heuristics about what might be worth pointing to).

Notice that this step of further constraining a pointing gesture also allows for a uniform treatment of **this** and **do this (that and do that)**. A preposed **this/that** may be the object of **do** or of some other verb, but the listener will not know which, until s/he reaches the verb itself, as in Example 9. Considering actions as properties of their respective events, the listener should be able to coerce **that** to be some appropriate facet of the discourse segment (or to some entity within that segment - as I will discuss next) that can be said or done.⁸

Example 9

Gladys told Sam last night that Fred was a complete jerk.

a. Anyway, **that's** what Fred believes that Gladys said.

b. Anyway, **that's** what Fred believes that Gladys did.⁹

On the other hand, what appears to be an additional ambiguity in resolving **this/that** may not be one at all. That is, a listener who is asked what a given **this/that** refers_m to must describe the representation that s/he has created. This act of description is subject to a lot of variability. For example, given a segment in which a statement A is supported by several pieces of evidence {B,C,D}, the listener might just describe A (the top level of the representation) or s/he might verbalize the whole representation.

As with anaphoric pronouns, when a deictic pronoun specifies an NP-evoked discourse entity, it must actually be part of its corresponding discourse segment interpretation. The interesting thing is that the same holds for deictic NPs, distinguishing them from anaphoric definite NPs, which can easily refer_m to things associated in some way with an existing entity, as in

Example 10

John and Mary decided to go on a picnic.
While they remembered most things,
they forgot to put the picnic supplies in the cooler.
So when they got to the park, **the beer** was warm.

By contrast, a similar example with a demonstrative NP sounds definitely odd --

Example 11

John and Mary decided to go on a picnic.
While they remembered most things,
they forgot to put the picnic supplies in the cooler.
#So when they got to the park, **that beer** was warm.

Another example illustrates this in another way: given that both anaphoric reference and deictic reference are possible in a particular context, an anaphoric NP and a deictic NP will be interpreted differently, even if in all other ways the NPs are the same. The anaphoric NP may refer_m to something associated with the current focus, while the deictic NP must point to something already explicitly included there. For example,

Example 12

- a. Some files are superfiles.
- b. To screw up some one's directory, look at **the files**.
- c. If one of them is a superfile,

Example 13

- a. Some files are superfiles.
- b. To screw up some one's directory, look at **those files**.
- c. They will tell you which of his files is absolutely vital to him.

In Example 12, **the files** is anaphoric, specifying the files in that person's directory, the entity currently in focus. In Example 13, **those files** is deictic, pointing to the files that are superfiles, i.e., to a discourse entity explicitly in the interpretation of the just current discourse segment.

Now, nothing in the process so far described distinguishes **this** and **that**. This is because with respect to discourse segment reference_m, it is rarely the case that the two cannot be used interchangeably.¹⁰ Thus it must be the case that this "psychological distance" feature of the deictic only comes into play after the referent_m is found. This does not imply though that **this** and **that** cannot have different effects on the discourse: in Sidner's 1982 theory [17] and in Schuster's theory of reference to actions and events [16], **this** and **that** are also distinguished by their effect (or lack thereof) on the **discourse focus**. This is compatible with it being a side effect of judging the speaker's "distance" from the referent_m, that the listener's beliefs about their shared discourse focus are revised.

To summarize, in Section 2, I argued for the existence of a second referring process associated with discourse segments *per se* rather than what they describe. In this section, I have argued for it having the features of pointing to the representation of a discourse segment on the right frontier, followed by further refinement to a property of the segment or an entity within its interpretation.

Here I want to argue for the proposed process having one additional feature. I have separated it out because it is not essential to the above arguments. However, it does permit an account of the common pattern of reference illustrated in Examples 1, 2, 14 and 15.

Example 1

It's always been presumed that when the glaciers receded, the area got very hot. The Folsom men couldn't adapt, and they died out. **That's** what is supposed to have happened. **It's** the textbook dogma. But **it's** wrong. They were human and smart. They adapted their weapons and culture, and they survived.

Example 2

The tools come from the development of new types of computing devices. Just as we thought of intelligence in terms of servomechanism in the 1950s, and in terms of sequential computers in the sixties and seventies, we are now beginning to think in terms of parallel computers, in which tens of thousands of processors work together. **This** is not a deep, philosophical shift, but **it** is of great practical importance, since it is now possible to study large emergent systems experimentally. [[6], p.176]

Example 14

I don't think **this** can be taken seriously either. **It** would mean in effect that we had learned nothing at all from the evaluation, and anyway we can't afford the resources **it** would entail.

Example 15

The Texas attorney general said that the McDonald's announcement represented "a calculated effort to make the public think that they were doing **this** out of the goodness of their heart when, in fact, they were doing **it** because of pressure from our office. [Philadelphia Inquirer, 13 June 1986]

Suppose one assumes that the ability to specify something via an anaphoric pronoun is a sufficient criterion for "discourse entity-hood". Then I would claim that whether or not a discourse segment referent_m is initially created as a discourse entity, once the speaker has successfully referred to it via **this/that**, it must now have the status of a discourse entity since it can be referenced via the anaphoric pronoun **it**.¹¹

Note that I do not mean to imply that one cannot refer deictically to the same thing more than once -- one clearly can, for example

Example 16

They wouldn't hear to my giving up my career in New York. **That** was where I belonged. **That** was where I had to be to do my work. [Peter Taylor, *A Summons to Memphis*, p.68]

Example 17

By this time of course I accepted Holly's doctrine that our old people must be not merely forgiven all their injustices and unconscious cruelties in their roles as parents but that any selfishness on their parts had actually been required of them if they were to remain whole human beings and not become merely guardian robots of the young. **This** was something to be remembered, not forgotten. **This** was something to be accepted and even welcomed, not forgotten or forgiven.

But of the (admittedly few) "naturally occurring" instances of this phenomenon that I have so far found, the matrix clauses are strongly parallel -- comments on the same thing. Moreover, except in cases such as Example 17, where the second clause intensifies the predication expressed in the first, the two clauses could have been presented in either order, which does not appear to be the case in the deixis-anaphor pattern of reference.

4. SUMMARY

In this paper, I have proposed and argued for a process-based account of subsequent reference via deictic expressions. The account depends on **discourse segments** having their own mental reality, distinct from that of the entities described therein. As such, discourse segments play a direct role in this theory, as opposed to their indirect role in explaining, for example, how the referents of definite NPs are constrained. One consequence is it becomes as important to consider the representation of entire discourse segments and their features as it is to consider the representation of individual NPs and clauses.

ACKNOWLEDGMENTS

This work was partially supported by ARO grant DAA29-884-9-0027, NSF grant MCS-8219116-CER and DARPA grant N00014-85K-0018 to the University of Pennsylvania, and an Alvey grant to the Centre for Speech Technology Research, University of Edinburgh. It was done while the author was on sabbatical leave at the University of Edinburgh in Fall 1987 and at Medical Computer Science, Stanford University in Spring 1988. My thanks to Jerry Hobbs, Mark Steedman, James Allen and Ethel

Schuster for their helpful comments on many, many earlier versions of this paper.

REFERENCES

- [1] Allen, J. *Natural Language Understanding*. Menlo Park: Benjamin/Cummings Publ. Co., 1987.
- [2] Cohen, R. A Computational Theory of the Function of Clue Words in Argument Understanding. *Proc. COLING-84*, Stanford University, Stanford CA, July 1984, pp.251-258.
- [3] Crain, S. and Steedman, M. On not being led up the garden path: the use of context by the psychological parser. In *Natural Language Parsing*, D. Dowty, L. Karttunen & A. Zwicky (eds.), Cambridge: Cambridge Univ. Press, 1985.
- [4] Grosz, B. The Representation and Use of Focus in a System for Understanding Dialogs. In *Elements of Discourse Understanding*, A. Joshi, B. Webber & I. Sag (eds.), Cambridge: Cambridge Univ. Press, 1981. (Reprinted in *Readings in Natural Language Processing*, B. Grosz, K. Sparck Jones & B. Webber (eds.), Los Altos: Morgan Kaufmann Publ., 1986.)
- [5] Grosz, B. & Sidner, C. Attention, Intention and the Structure of Discourse. *Computational Linguistics*, 12(3), July-Sept. 1986, pp.175-204.
- [6] Hillis, W.D. Intelligence as an Emergent Behavior, *Daedalus*, Winter 1988, pp.175-190.
- [7] Hirschberg, J. & Litman, D. Now Let's Talk about Now: Identifying Cue Phrases Intonationally. *Proc. 25th Annual Meeting, Assoc. for Comp. Ling.*, Stanford Univ. Stanford CA, July 1987.
- [8] Hobbs, J. , Stickel, M., Martin, P. and Edwards, D. Interpretation as Abduction. *Proc. 26th Annual Meeting, Assoc. for Comp. Ling.*, SUNY Buffalo, Buffalo NY, June 1988.
- [9] Karttunen, L. Discourse Referents. In *Syntax and Semantics*, Volume 7, J. McCawley (ed.), New York: Academic Press, 1976.
- [10] Miller, G. Problems in the Theory of Demonstrative Reference. In *Speech, Place and Action*, R. Jarvella & W. Klein (eds.), New York: Wiley, 1982.

- [11] Moens, M. and Steedman, M. Temporal Ontology and Temporal Reference. **Computational Linguistics**, to appear Summer 1988.
- [12] Nakhimovsky, A. Aspect, Aspectual Class and the Temporal Structure of Narrative. **Computational Linguistics**, to appear Summer 1988.
- [13] Polanyi, L. The Linguistic Discourse Model: Towards a formal theory of discourse structure. TR-6409. BBN Laboratories Incorp., Cambridge MA, November 1986.
- [14] Quine, W. The Inscrutability of Reference. In *Semantics: An Interdisciplinary Reader*, D. Steinberg & L. Jacobovits (eds.), Cambridge: Cambridge University Press, 1971. pp.142-154.
- [15] Reichman, R. **Getting Computers to Talk like You and Me**. Cambridge MA: MIT Press, 1985.
- [16] Schuster, E. Pronominal Reference to Events and Actions: Evidence from Naturally-occurring data. MS-CIS-88-13, Computer & Information Science, Univ. of Pennsylvania, February 1988.
- [17] Sidner, C. Focusing in the Comprehension of Definite Anaphora. In **Computational Models of Discourse**, M. Brady & R. Berwick (eds.), Cambridge MA: MIT Press, 1982, pp.267-330.
- [18] Webber, B. So What can we Talk about Now? In **Computational Models of Discourse**, M. Brady & R. Berwick (eds.), Cambridge MA: MIT Press, 1982, pp.331-371.

¹ The five texts are (1) Peter Taylor's novel, *Summons to Memphis*, Ballentine Books, 1986 (pp.1-21); (2) W.D. Hillis' essay, "Intelligence as an Emergent Behavior", *Daedalus*, Winter 1988, pp.175-189; (3) an editorial from *The Guardian*, 15 December 1987; (4) John Ryle's review of a set of books on drug use, "Kinds of Control", *TLS*, 23-29 October 1987, pp.1163-1164; (5) Phil Williams' review of a set of books on disarmament, "New threats, new uncertainties", *TLS*, 20-26 November 1987, p.1270. All instances of pronominal reference_m using **it**, **this** and **that** were tabulated.

I specifically used written (primarily objective) expositions rather than spoken texts in order to avoid

the common use of **this/that** in first-person accounts to refer to the outside world.

² that is, ignoring all syncategorematic uses of it (as in "It is possible that John is here")

³ As I shall argue at the end of Section 3, the ability to refer to something anaphorically might be a sufficient, though perhaps not a necessary criterion for "entity-hood".

⁴ If the example were "**That's** all I know about it", **that** would be taken as referring to the description of House B, not the discourse segment associated with the clause "I heard all this from a friend, who saw the house yesterday". (Call this later segment DS-h.) However, this need not invalidate my claim about the accessibility of discourse segments since DS-h can be understood as a parenthetical, which are treated differently than non-parentheticals in theories of discourse - cf. [GS85]. While a parenthetical may itself contain a deictic pointer to a discourse segment on the right frontier, it doesn't influence the frontier. Thus **that** still has the same discourse segments accessible as it would without the parenthetical. Another example of discourse deixis from a parenthetical is this variation of Example 5.

...it should be possible to identify certain functions as being unnecessary for thought by studying patients whose cognitive abilities are unaffected by locally confined damage to the brain. For example, binocular stereo fusion is known to take place in a specific area of the cortex near the back of the head. (**This** was discovered about 10 years ago). Patients with damage to this area of the cortex have visual handicaps but show no obvious impairment in their ability to think.

⁵ To get further data on this, I ran an informal "discourse completion" experiment, modelled on the above lines, presenting a short, multi-sentence text which I judged as having several segments on the right frontier at the point of the last sentence. As above, I asked subjects to complete a next sentence beginning "**That...**"

<The subject here is legends of the formation of the Grand Canyon>

<What follows is the second paragraph of the given text>

"Another legend tells of a great chief who could not cease from mourning the death of his beloved wife. Finally the gods offered to take him to visit his wife

so that he could see she was contented in the happy hunting ground. In exchange, he was to stop grieving when he returned to the land of the living. **That...**"

I also asked subjects to paraphrase what they wrote, to see explicitly what they took **that** to specify. The responses I got showed them taking it to specify either the chief's action (expressed in the previous, single sentence segment) or the whole "bargain" (expressed in the segment comprising both previous clauses). While this particular experiment was only informal and suggestive, well-controlled versions should be able to produce harder results.

⁶ Presumably A_Junior will have enough context to resolve this more precisely, or he will be smart enough to ask.

⁷ Of the 69 clausally-referring instances of **this** and **that** pronouns, 51 (~70%) were in subject position in standard SVO clauses (7 instances of **that** and 44, of **this**), 17 played some other role within their matrix clause, and 1 was a preposed adverbial ("after that"). Hence ~75% were first NPs.

⁸ This does not say which of those actions will be picked out. See [Schus88] for a discussion of the choice of event/action referents of pronouns.

⁹ It is possible to construct quite acceptable examples in which a preposed **that** functions as the object of both **do** and some other verb -- for example "Several universities have made computer science a separate school. But **that** is not necessarily what we want or could even do." The conjunction of two forms usually means that at some level, both forms are taken as being the same.

¹⁰ **That** is because with respect to discourse segment reference_m, it is rarely the case that the two cannot be used interchangeably!

¹¹ If one assumes that a discourse segment referent_m is also a discourse entity *ab ovo*, as it were, then this pattern might simply be interpreted as such an entity coming into focus as a result of the deictic reference. As I noted earlier, there is not enough evidence to argue either way yet, nor is it clear that the two accounts would have vastly different consequences anyway.