Identifying Subjective Characters in Narrative

Janyce M. Wiebe
Department of Computer Science
University of Toronto
Toronto, Canada M5S 1A4
wiebe@ai.toronto.edu

ABSTRACT

Part of understanding fictional narrative text is determining for each sentence whether it takes some character's point of view and, if it does, identifying the character whose point of view is taken. This paper presents part of an algorithm for performing the latter. When faced with a sentence that takes a character's point of view, the reader has to decide whether that character is a previously mentioned character or one mentioned in the sentence. We give particular consideration to sentences about private states, such as seeing and wanting, for which both possibilities exist. Our algorithm is based on regularities in the ways that texts initiate, continue, and resume a character's point of view, found during extensive examinations of published novels and short stories.

1. INTRODUCTION.

Part of understanding fictional narrative text is determining for each sentence whether it takes some character's *point of view* and, if it does, identifying the character whose point of view is taken. This paper addresses the latter. We show how structural regularities of third-person fictional narrative text can be exploited to perform this task.

2. BACKGROUND.

Our approach to point of view is based on Ann Banfield's (1982) theory of narrative sentences. Banfield characterizes the sentences of narration as *objective* or *subjective*. Objective sentences are those that objectively narrate events. Subjective sentences are those that present the consciousness of an experiencing character within the story. They express characters' evaluations, emotions, judgments, uncertainties, beliefs, and other attitudes and affects. Kinds of subjective sentences include those that portray a character's thought or perception (*represented thought* or *represented perception*; Banfield 1982), and those that report a character's *private state* such as seeing, wanting, or feeling ill—that is, some perceptual, psychological, or experiential

state that is not open to objective observation or verification.[1] We call the character whose consciousness is presented by such a sentence the *subjective character* of that sentence.

The following passage illustrates different kinds of subjective sentences.

(1)
    1.1"What are you doing in here?" 1.2*Suddenly she [Zoe] was furious with him [Joe].*
    1.3"Spying, of course."
    1.4"Well of all dumb things! 1.5I thought you ran away." 1.6*Joe Bunch was awful.* [Oneal, *War Work*, p. 130]

Sentence (1.2) is a private-state report: It reports Zoe's private state of being furious with Joe. Sentence (1.6) is a represented thought: It presents Zoe's thought and it expresses her evaluation of Joe (that he is awful). Represented perceptions appear in the following passage:

(2)
    2.1*Certainly, Dennys thought, anything would be better than this horrible-smelling place full of horrible little people.*
    2.2*There was a brief whiff of fresh air.* 2.3*A glimpse of a night sky crusted with stars.* [L'Engle, *Many Waters*, p. 25]

Sentence (2.1) is Dennys's represented thought and (2.2) and (2.3) are Dennys's represented perceptions.

Third-person narrative text is characterized by shifts in point of view: a single text can have objective sentences as well as subjective sentences attributed to different characters.[2] To determine the current point of view, readers cannot consider sentences in isolation. In many cases, it is only in

---

[1] See Wiebe 1990 and the references therein for alternative categorizations of subjective sentences.

[2] There are different kinds of point of view in narrative. The kind addressed here is *psychological* as opposed to, for example, *spatial* or *temporal* (Uspensky 1973). In addition, we do not consider texts that take the point of view of an *overt narrator* (Chatman 1978).

context that a sentence is subjective or has a particular subjective character. Sentence (1.6) is such a sentence.

We have developed an algorithm for recognizing subjective sentences and identifying their subjective characters in third-person fictional narrative text. The algorithm is based on regularities, found during our examination of passages from over forty novels and short stories, in the ways that texts initiate, continue, and resume a character's point of view. The rules of the algorithm were checked, by hand, on over four hundred pages from seven novels. We were able to categorize most exceptions according to particular problems that remain to be addressed, such as the effect of certain spatial and temporal discontinuities on the psychological point of view. These classes of exceptions, together with complete descriptions of the current algorithm and its implementation, can be found in Wiebe 1990. A preliminary version of the algorithm was presented by Wiebe & Rapaport 1988.

3. IDENTIFYING THE SUBJECTIVE CHARACTER.

3.1. Introduction. Black, Turner, & Bower (1979) claim the following result of their empirical investigation of point of view in narrative: "Merely making a character the subject of the narrative statement sufficed to establish his as the dominant point of view" (p. 187). However, their stimulus materials were short, artificially constructed narratives. For extended texts (i.e., novels and short stories), this simple rule is inadequate.

The subjective character of a subjective sentence *is* sometimes identifiable from the sentence itself. This is the case if the sentence has a narrative parenthetical (such as 'Dennys thought' in (2.1)); it might also be the case if the sentence denotes a private state, but not always. If the subjective character is not identifiable from the sentence, it is often one of two previously mentioned characters: the subjective character of the previous subjective sentence or the actor of an action denoted by a previous objective sentence. The availability of these characters is conditional; obviously, for example, the former is available only if a subjective sentence has previously appeared. The subjective character also may not be identifiable when the sentence appears.

Thus, identifying the subjective character involves determining whether it can be identified

from the sentence itself, or, if not, determining whether it can be identified from the previous context. In addition, as the text is processed, information that might be needed to identify later subjective characters must be gathered.

3.2. Identifying the Subjective Character from the Previous Context. A subjective character who is not identifiable from the sentence itself (see Section 3.3, below) is most often the subjective character of the last subjective sentence. In this case, the current sentence continues a character's point of view (if the previous sentence was also subjective) or resumes one (if objective sentences have appeared since the last subjective sentence). Sentences (2.2) and (2.3) illustrate the former, and sentence (1.6) illustrates the latter; it has the same subjective character as (1.2), the last subjective sentence, and is separated from (1.2) by objective sentences (1.3)–(1.5).

Less commonly, the subjective character is the actor of an action denoted by a previous objective sentence. Since this character need not be the subjective character of the last subjective sentence, this is a way to initiate a new point of view. Following is an example:[3]

(3)
3.1Jake felt sour. 3.2He wished again that circumstances hadn't prompted him to come back. 3.3He had already spent one full night on horseback, 3.4and now the boys were expecting him to spend another, all on account of a bunch of livestock he had no interest in in the first place. 3.5"I don't know as I'm coming," he said. 3.6"I just got here. 3.7If I'd known you boys did nothing but chase horses around all night, I don't know that I would have come."
3.8"Why, Jake, you lazy bean," Augustus said, 3.9and walked off. 3.10*Jake had a stubborn streak in him,* 3.11*and once it was activated even Call could seldom do much with him.* [McMurtry, *Lonesome Dove*, p. 162]

As this passage appears in the novel, (3.10) and (3.11) represent the point of view of Augustus, the

---

[3] Numbering within passages reflects the actual units of input given to the algorithm, which are not always sentences. E.g., the conjuncts of compound sentences are separate input units. Even so, an input unit can still have multiple clauses, each of which denotes an event. The algorithm uses heuristics to choose the most important of these for tracking point of view (most often, it chooses the event denoted by the main clause). Thus, when we refer to a sentence that denotes an action, for example, we actually mean an input unit of which the event chosen for consideration is an action.

actor of an action denoted by a previous objective sentence, (3.9). But the last subjective character is Jake, so Augustus's point of view is initiated, not merely resumed or continued.

In order to identify the subjective character in these situations, the algorithm keeps track of *expected subjective characters*, characters who are likely to become the subjective character. Often, the last subjective character is one of these. However, drastic spatial and temporal discontinuities can block the continuation or resumption of a character's psychological point of view. We consider one kind of discontinuity, a *scene break*, which is a break from one parallel story-line to another (Almeida 1987). Thus, the last subjective character is an expected subjective character only if a subjective sentence has appeared in the current scene (see Wiebe 1990 for examples).[4]

The usual situation for a shift in point of view to an actor is one in which the action of a previous subjective character is focused by the text—specifically, when the objective sentence denoting the action appeared earlier in the current paragraph and the entire paragraph so far has been objective.[5] In these situations, the actor of the last such sentence, the *last active character*, is an expected subjective character.

If the subjective character is not identifiable from the sentence, then the algorithm chooses an expected subjective character, if there is one. If there isn't one, then the algorithm cannot identify the subjective character at this point in the text (see Wiebe & Rapaport 1988 and Wiebe 1990 for examples in which the reader must read further in the text in order to identify the subjective character). In the case of competition between expected subjective characters, the last active character is the preferred choice, since he or she is more highly focused by the text than the last subjective character is. The last subjective character is chosen only if the sentence is about the last active character.[6] The choice of the last *subjective* character is incorrect if the sentence is the last *active* character's self-reflection or self-perception; this

---

[4] The algorithm cannot detect scene breaks; rather, it must be informed when one occurs.

[5] In addition, the action must be one that is actually performed in the current scene; thus, the main verb phrase of the sentence denoting it must pass certain linguistic tests, for example, that it not contain a habitual adverbial such as 'usually' or a modal adverbial such as 'probably'.

[6] That is, if the last active character is the actor or experiencer of the event chosen for consideration.

heuristic relies on the relative infrequency of subjective sentences about oneself.

3.3. Identifying the Subjective Character from the Sentence.
Most commonly in third-person narrative text, subjective sentences continue or resume characters' points of view, rather than initiate new ones. The previous section discussed one way, illustrated by passage (3), to initiate a new point of view. In situations like those, the subjective character is determined from the previous context. A subjective sentence is also able to initiate a new point of view if the subjective character is identifiable from the sentence itself. A simple example is a sentence with a narrative parenthetical, such as (2.1)—the subjective character is always the subject of the parenthetical.

The subjective character may also be identifiable from a sentence denoting a private state (a *private-state sentence*); in particular, it may be the experiencer of the private state denoted. Thus, as noted by Doležel 1973, Cohn 1978, and Banfield 1982, a private-state sentence is another way to initiate a new point of view.[7] An example occurs in the following:

(4)
4.1"Drown me?" Augustus said. 4.2"Why if anybody had tried it, those girls would have clawed them to shreds." 4.3He knew Call was mad, 4.4but wasn't much inclined to humor him. 4.5It was his dinner table as much as Call's, 4.6and if Call didn't like the conversation he could go to bed.
4.7Call knew there was no point in arguing. 4.8That was what Augustus wanted: argument. 4.9He didn't really care what the question was, 4.10and it made no great difference to him which side he was on. 4.11He just plain loved to argue. [McMurtry, *Lonesome Dove*, p. 16]

Sentences (4.3)–(4.6) are Augustus's subjective sentences and (4.7)–(4.11) are Call's. So, (4.7) initiates a new point of view. It is a private-state sentence and the subjective character, Call, is the experiencer of the private state denoted. Passage (4) shows, however, that the subjective character of a private-state sentence need not be the experiencer. In (4.6), for example, "Call didn't like the conversation" denotes a private state (liking), but the subjective character of the sentence is

---

[7] If a subjective sentence denotes a *private-state action* such as sighing or looking, the subjective character might also be identifiable from the sentence itself. See Wiebe 1990 for the algorithm's treatment of these kinds of sentences.

Augustus, not Call. In the following subsections, we will consider factors that can indicate that the subjective character of a private-state sentence is not the experiencer.

3.3.1. Textual Continuity.

Point of view does not typically shift from one character to another without a paragraph break.[8] Thus, the lack of a paragraph break suggests that a shift has not occurred. In the case of a private-state sentence following a subjective sentence, where the experiencer is different from the last subjective character, the lack of a paragraph break between them suggests that the private-state sentence is *not* the experiencer's subjective sentence; if it were, then a shift would have occurred without a paragraph break. So, when a private-state sentence immediately follows a subjective sentence without a paragraph break (this is called the *continuing-subjective* text situation), the algorithm identifies the subjective character to be the last subjective character, not the experiencer of the private state.[9]

3.3.2. Subjective Elements.

Another factor that suggests that a private-state sentence should be attributed to an expected subjective character rather than to the experiencer is the appearance of a *subjective element*.

There are many linguistic elements (syntactic, lexical, and morphological) that can indicate that a sentence is subjective. Examples are evaluative terms such as 'the old bag' (Banfield 1982), and evidentials such as 'evidently' and 'apparently' (Doležel 1973). Although some of these elements always indicate that a sentence is subjective, others only potentially do so. A *subjective element* is a linguistic element that actually indicates that a sentence is subjective in the context of use.[10]

Before we address the effect of subjective elements on identifying the subjective character of a private-state sentence, we need to consider interpretations of private-state sentences. There

are two subjective interpretations of private-state sentences: represented thoughts and private-state reports.[11] Interpreted to be a represented thought, a private-state sentence is someone's thought *about* a character's private state. For example, (4.6) is Augustus's represented thought that is about Call's private state (Call's not liking something). A private-state sentence that is interpreted to be a private-state report, on the other hand, is a report *of* a character's private state; it is not someone's thought about it. Examples are (4.3) and (4.7).[12]

The subjective character of a private-state *report* is always the experiencer. Thus, the subjective character is identifiable from the sentence itself. On the other hand, the subjective character of a represented thought cannot be identified from the sentence itself, but depends on the context (unless, of course, a narrative parenthetical appears). This is so regardless of whether the subjective character is referred to in the sentence. For example, the subjective character of the following represented thought is the referent of 'she' (as the sentence appears in the short story):

Why couldn't she? [Mansfield, "The Garden Party"]

but the subjective character of the following (as the sentence appears in the short story) is not referred to in the sentence:

What nice eyes he had, small, but such a dark blue! [Mansfield, "The Garden Party"]

Thus, if something indicates that a private-state sentence is a represented thought rather than a private-state report, then the subjective character should be identified to be an expected subjective character (as discussed in Section 3.2).

Subjective elements do this, as in the following passage; at the start of the passage, Sandy and Dennys are (collectively) the last subjective character:[13]

---

[8] This observation is predicted by Nakhimovsky's work on the discourse structure of narrative text, in which it is suggested that paragraph breaks accompany discontinuities (Nakhimovsky & Rapaport 1988; Nakhimovsky 1988).

[9] Actually, the algorithm allows for a *broadening* or *narrowing* of point of view upon a private-state sentence. This can occur because, as shown by Banfield 1982, a subjective sentence can be attributed to a set of characters.

[10] We borrow this term from Banfield 1982, but redefine it; Banfield uses it to refer to linguistic elements that *always* indicate that a sentence is subjective.

[11] There is also an objective interpretation (see Wiebe 1990).

[12] Cohn 1978 does not acknowledge this ambiguity: She says that represented thought can be distinguished from private-state reports by "the absence of mental verbs" in the former (p. 104).

[13] According to the rules of the algorithm, the event of (5.1) chosen for consideration is the private state denoted by the subordinated clause.

(5)
  [5.1]Japheth, *evidently* realizing that they were no longer behind him, turned around [5.2]and jogged back toward them, seemingly cool and unwinded. [L'Engle, *Many Waters*, p. 24]

The subjective element 'evidently' in (5.1) indicates that the sentence is *not* a private-state report. That is, (5.1) is not a report that Japheth realizes that they are no longer behind him. Instead, Sandy and Dennys (the subjective character) ascribe this private state to him.

However, *subordinated* subjective elements, those *within* the scope of the private-state term, can appear in private-state reports.[14] Thus, they cannot be used to distinguish private-state reports from represented thoughts (and so cannot be used to identify the subjective character). For example:

(6)
  [6.1]Ugh! she [the girl] thought. [6.2]How could the poor thing have married him in the first place?
  [6.3]Johnnie Martin could not believe that he was seeing that *old bag's* black eyes sparkling with disgust and unsheathed contempt at him. [Caldwell, *No One Hears But Him*, pp. 98–99]

Sentence (6.3) is a private-state report and the subjective character is the experiencer (Johnnie Martin); this is so even though (6.3) contains the subjective element 'old bag' and even though there is an expected subjective character (the girl) when it is encountered. Because 'old bag' appears within the scope of the private-state term 'believe', it is not considered in identifying the subjective character. On the other hand, the subjective element 'evidently' in (5.1) is not in the scope of 'realizing' (i.e., it is *non-subordinated*), so it can be used to identify the subjective character.

If a private-state sentence does not have a non-subordinated subjective element and does not appear in the continuing-subjective situation, then the algorithm identifies the expected subjective character to be the experiencer.

4. SUMMARY. Following is the rule of the algorithm for identifying the subjective character of a subjective sentence ('SC' stands for the subjective character of the current sentence):[15]

---

[14] This is one reason why we define private-state reports to be subjective.

[15] Provisions for the following have been omitted for brevity: broadening and narrowing of point of view, sentences denoting private-state actions, and private-state sentences that can be objective

To identify SC:
If the sentence contains a narrative parenthetical then
    SC is the subject of the parenthetical
else if the sentence is a private-state sentence then
    if it has a non-subordinated subjective element
    or the text situation is continuing-subjective then
        SC is identified from the previous context
    else SC is the experiencer
    end if
else SC is identified from the previous context
end if

To identify SC from the previous context:
If there are two expected subjective characters then
    if the sentence is about the last active character then
        SC is the last subjective character
    else SC is the last active character
    end if
else if there is an expected subjective character then
    SC is the expected subjective character
else SC is unidentified
end if

5. CONCLUSION. We are extending this work along two avenues. First, we are developing psychological experiments to test whether the regularities on which the algorithm is based influence the reader's recognition of subjective sentences and identification of subjective characters. Second, we are extending the algorithm to make connections with work on focus of attention and discourse structure (such as Grosz & Sidner 1986); in particular, we are investigating how resolving anaphora and tracking the current point of view are related (Stark 1987, Hewitt 1988). An important direction for future research is reasoning about the plausibility of a suggested interpretation, that is, whether it is plausible that the content of a subjective sentence is a particular character's thought or perception.

We have presented part of an algorithm for identifying subjective characters that is based on regularities in the ways that texts initiate, resume, and continue a character's point of view. When faced with a subjective sentence, the reader has to decide whether the subjective character can be identified from the sentence itself. It can be if the sentence is a private-state report, but not if it is a represented thought; we give rules for distinguishing between these two interpretations of private-state sentences that involve the textual situation and the appearance of subjective elements in the sentence. A subjective character who is not

---

(see Wiebe 1990).

identifiable from the sentence itself is often one of two previously-mentioned characters; we identify those characters and specify when one of them should be chosen to be the subjective character. Our findings are based on extensive examinations of natural narrative.

REFERENCES

Almeida, M. J. (1987), *Reasoning about the Temporal Structure of Narratives*, Ph.D. dissertation, published as *Technical Report 87–10* (Buffalo: SUNY Buffalo Department of Computer Science).

Banfield, A. (1982), *Unspeakable Sentences: Narration and Representation in the Language of Fiction* (Boston: Routledge & Kegan Paul).

Black, J. B.; Turner, T. J.; & Bower, G. H. (1979), "Point of View in Narrative Comprehension, Memory, and Production," *Journal of Verbal Learning and Verbal Behavior* 18: 187-198.

Caldwell, Taylor (1966), *No One Hears But Him* (Garden City, NY: Doubleday).

Chatman, Seymour (1978), *Story and Discourse: Narrative Structure in Fiction and Film* (Ithaca, NY: Cornell University Press).

Cohn, D. (1978), *Transparent Minds: Narrative Modes for Representing Consciousness in Fiction* (Princeton: Princeton University Press).

Doležel, L. (1973), *Narrative Modes in Czech Literature* (Toronto: University of Toronto Press).

Grosz, B. J. & Sidner, C. L. (1986), "Attention, Intentions, and the Structure of Discourse," *Computational Linguistics*, 12: 175-204.

Hewitt, L. (1988), "Reduced Anaphor in Subjective Contexts," unpublished manuscript, SUNY Buffalo Graduate Group in Cognitive Science.

L'Engle, M. (1986), *Many Waters* (New York, Dell Publishing).

McMurtry, L. (1985), *Lonesome Dove* (New York: Simon & Schuster).

Nakhimovsky, A. (1988), "Aspect, Aspectual Class, and the Temporal Structure of Narrative," *Computational Linguistics* 14 (2): 29-43.

Nakhimovsky, A. & Rapaport, W. J. (1988), "Discontinuities in Narratives," *Proceedings of the 12th International Conference on Computational Linguistics (COLING-88, Budapest)* (Morristown, NJ: Assoc. for Computational Linguistics): 465-470.

Oneal, Z. (1971), *War Work* (New York: Viking Press).

Stark, H. A. (1987), "Keeping Track of Characters in Narrative," Ph.D. dissertation, Dept. of Computer Science, Stanford University.

Uspensky, B. (1973), *A Poetics of Composition* (Berkeley: University of California Press).

Wiebe, J. M., & Rapaport, W. J. (1988), "A Computational Theory of Perspective and Reference in Narrative," *Proceedings of the 26th Annual Meeting of the Assoc. for Computational Linguistics (ACL-88; SUNY at Buffalo)* (Morristown, NJ: Assoc. for Computational Linguistics): 131-138.

Wiebe, J. M. (1990), *Recognizing Subjective Sentences: A Computational Investigation of Narrative Text*, Ph.D. dissertation, published as a technical report (Buffalo: SUNY Buffalo Dept. of Computer Science).