

Department of Computer Science

Departmental Papers

---

University of Texas at El Paso

Year 2001

---

**"Conversational" Dialogues in Direct-  
Manipulation Interfaces**

David G. Novick, University of Texas at El Paso

This paper was published as: Novick, D., (2001). "Conversational" Dialogues in Direct-Manipulation Interfaces, *Proceedings of the 2nd IJCAI Workshop on Knowledge and Reasoning in Practical Dialogue Systems*, Seattle, WA, August 5, 2001

This paper is posted at DigitalCommons@UTEP  
[http://digitalcommons.utep.edu/cs\\_papers/8](http://digitalcommons.utep.edu/cs_papers/8)

# "Conversational" Dialogues in Direct-Manipulation Interfaces

**David G. Novick**

Department of Computer Science  
University of Texas at El Paso

## **Abstract**

This paper reports ongoing research in extending direct-manipulation interfaces by incorporating, via the direct-manipulation modality itself, interaction techniques that add kinds of language features associated with spoken conversation. The paper proposes means of implementing ways for a user of a direct-manipulation system to define new kinds of relations among objects in the interface.

## **1 INTRODUCTION**

Can traditional GUI interfaces can be extended to include some of the more dynamic functions associated with spoken conversation? This paper suggests that direct-manipulation (DM) interfaces can be extended by incorporating, via the DM modality itself, interaction techniques that add certain language features associated with spoken conversation.

Clark and his colleagues [Clark and Marshall, 1981; Clark and Wilkes-Gibbs, 1986] reported that, in human-human conversations, the conversants formed mutual beliefs about referents, and that this mutuality extended to the way in which expressions about these referents were generated and understood. Other work [e.g., Lambert and Carberry, 1992] had explored negotiation subdialogues about domain tasks. What Clark et al. observed, though, was a kind of negotiation about the means through which concepts in the conversation were grounded. The conversants were able to create, in effect, new relationships that applied to things within the domain of their conversation. This led to use referring expressions that were shorter and more direct.

This kind of joint action to create new relationships for referring to things of interest is problematic for DM, particularly as implemented via WIMP interfaces. The principles of DM were originally articulated by Shneiderman [1983]:

- Continuous representation of the objects of interest
- Physical actions on objects vs. complex syntax
- Fast, incremental and reversible operations with an immediately apparent effect on the objects of interest
- Layered or spiral approach to learning

As implemented in typical WIMP interfaces, the simplicity of physical action limits relations to those that have already been defined by the system for the user. For example, a user can indicated that an object belongs to a class of objects by dragging the object into a folder. But what about relations other than membership in a class? How, other than through paraverbal means such as dialogue boxes, could a user define a new relation? Such negotiation of relations and of meaning are easy enough in conversation to occur without notice, yet virtually impossible in a DM interface.

## **2 MORE FULLY REALIZED DM INTERFACES**

Beaudoin-Lafon's recent work [2000] on interaction instruments showed that Shneiderman's principles of DM have been far from fully realized. Going beyond the current generation of WIMP interfaces, Beaudoin-Lafon proposed creating, in DM interfaces, a kind of instrument as mediator of actions into commands, resulting in reactions in the interface. The interaction instrument is a kind of transducer that transforms users' actions into commands affecting domain objects. For example, it is possible to create an interaction instrument for search-and-replace functions that (a) do not use buttons and (b) enable a simultaneous view of multiple instances of the searched-for string as it occurs in the text. This work demonstrated that following Shneiderman's principles could lead to post-WIMP interfaces that significantly increased functionality provided to users.

Yet even in this extended view DM is best at things such as pointing at objects (i.e., direct reference) and moving objects (i.e., applying predefined relationships to objects). One might surmise that there is something inherently limiting about continuous representation of objects of interest, or about adhering to physical actions in place of complex syntax. So, inspired by Beaudoin-Lafon's effort, I began to explore what it would mean to express Clarkian notions of negotiation of meaning and relationships while still adhering to Shneiderman's principles. Indeed, the point of adhering to these principles is that the directness of DM interfaces might be enabled for these fairly subtle functions of conversation.

## **3 AN EXAMPLE: WRITING AN ABSTRACT**

Here is the canonical example that motivates the negotiation of new relations in a DM interface. Consider the an author who would like to create a first-draft abstract of an article by assembling the first sentence of each paragraph of the article into a new, summary paragraph. In current DM interfaces, this would typically involve, for each paragraph of the article, selecting the first sentence of the paragraph, copying it, and pasting the sentence into a new text field. Or, the author could possibly perform a multiple selection by carefully selecting every first sentence, copying all of them at once, and then pasting all of them into the new text field. The first approach is tedious; the second approach is both tedious and risky, because a slip when doing the multiple selection might lead to the loss of much of the multiple-selection work.

If the author could converse with the text system, he or she might instead negotiate the definition of a new relation for the interface, something like "first sentence of a paragraph", and then apply this new relation to the article to select the text. Note that building the "first sentence of a paragraph" relation into the interface is not a general solution to this sort of problem, as the author might want to use some other relation later on. Building in every possible relation in advance would seem (a) to be impractical and (b) to lead to reference confusion with respect to relations. So the solution must lie in providing means for users and systems to negotiate new relations.

## **4 BASES FOR A SOLUTION**

The proposed solution is based on the idea of situated acts, which are conversational acts abstracted from modality [Novick and Perez-Quinones, 1998]. Situated acts grow out of

situated action [Suchman, 1987], except that they operate entirely at the level of domain acts rather than at the level of actions in the interface. This meant transforming the problem of creating new relations in a DM interface to a more abstract problem of creating new relations independent of modality and then “specializing” the solution into the DM interface, if possible. So how could a “conversational” function like negotiation of the definition of a relation be incorporated in a DM interface? At the more abstract, situated-acts level, analysis suggests that two things are needed to do this: (1) a means to perform ordered selection and (2) a common set of concepts for expressing relations. The property of ordered selection is necessary because relations are not necessarily symmetric. For example, “A is to the left of B” is not equivalent to “B is to the left of A.” The common set of concepts is needed in order to build new meaning from knowledge that can be assumed to be already mutual. This is clearly the case in natural-language dialogues; this kind of a priori common ground is no less necessary in a DM interface.

## 5 APPLYING THE SOLUTION TO THE EXAMPLE

There appear to be numerous ways of implementing ordered selection and a common vocabulary. To illustrate the basic ideas of the approach, I developed a conceptual prototype that implements these ideas in concrete terms for the “writing an abstract” example. In the conceptual prototype, ordered selection is implemented via two, different-colored cursors, and a common set of concepts for describing possible relations is implemented as a built-in set of terms such as letter, word, sentence, paragraph and ordinals.

The interface's design enables the user to perform the following steps:

- Select a sentence with one cursor
- Select the sentence's paragraph with a second cursor
- Observe that the relation between the two selections is “first sentence of a paragraph”
- Apply the relation to the entire document

These steps are illustrated in Figures 1 through 4. Figure 1 shows the text field, a “cursor palette” and a schema for a relation. The text field contains the original text to be abstracted. The cursor palette enable the user to choose which of two cursors they will use in the text field. In the conceptual prototype the cursors are red and blue. In this paper, they are rendered as light gray and dark gray.

Figure 2 shows the state of the interface after the user has used the red (light gray) cursor to select the first paragraph in the text field. The start of the highlighted text appears in the field representing the first argument of the relation; this text is flagged by a red P, signifying that the system understands that a paragraph has been selected.

Figure 3 shows the state of the interface after the user has used the blue (dark gray) cursor to select the first sentence of the first paragraph. The start of the highlighted text appears in the field representing the second argument of the relation; this text is flagged by a red S, signifying that the system understands that a sentence has been selected. Additionally, the interface displays a name describing its understanding of the relation

between the two selected things. In this case, the name is “First Sentence.” The name reflects the interface’s understanding of the ordinal nature of the defined relation.

Finally, Figure 4 shows the text field after the user has applied the newly defined “First Sentence” relation to the entire text field. All of the text is highlighted in light gray, and the first sentence of every paragraph is highlighted in dark gray. It remains for the user to copy the highlighted first sentences and paste them into a different text field to create the draft abstract.

## CONCLUSION

Through the use of ordered selection and a common set of concepts, the user and the system have reached a kind of understanding about the meaning of a relation that did not previously exist in the interface. In effect, the user is employing meta-speech-acts through DM interaction facilities provided in the interface. The relation-defining schema can be viewed as a kind of Beaudoin-Lafon interaction instrument that uses DM to produce meta-commands in the interface. This approach is entirely consistent with Shneiderman’s principles:

- The text, selections and relations have continuous representations.
- The user’s actions on the text are entirely physical; the verbal output comes from the system as a means of confirming that the system and user have a mutual understanding of the user’s actions.
- The available operations are fast, incremental and reversible, with an immediately apparent effect on the objects of interest. The user could as easily unselect text in this prototype as in more traditional text editors.
- The system’s feedback in the relation schema enables the user to explore the system’s concept space, thus providing for a layered or spiral approach to learning.

However, the approach as embodied in the proposed prototype is subject to a number of limitations. These include the fact that the nature of the possible relations is constrained by the limited conceptual vocabulary and, more significantly, the lack of negotiation in the process of reaching the “understanding” between user and system. Because the interface provides an interaction instrument, the user is, in effect, issuing a set of commands to the system. The system does not yet have the means (a) to signal uncertainty or confusion about the user’s intent (other than simply getting it wrong), or (b) to suggest alternate meanings or new relations of its own. Future work on this problem thus should include exploring (a) how to add new mutually understood concepts to the set of building blocks for the relations, (b) ways of negotiating rather than commanding, and (c) means to make these dialogues more symmetric. Another avenue of research involves exploring alternate ways of expressing ordered selection. And the user’s selections may be interpretable as more than one proposed relation; how can the user and system resolve this kind of ambiguity?

Taking a longer view, are there other kinds of “conversational” features that would be useful and practical to implement in direct-manipulation interfaces? And if this kind of meta-interaction became prevalent, what would be the effects on user? Just as human conversants use knowledge of the identity of other conversants to establish mutual

knowledge, so too may users and systems have to establish each other's identity to determine the kinds of dialogue acts that they can reasonably expect the other party to understand.

## **ACKNOWLEDGMENTS**

Sanjaya Das, Ganapathi Adimurthy, Isaac Hernandez, Asis Nayak and Karen Ward contributed to this research. This material is based upon work supported by the National Science Foundation under Grant No. 0080940.

## **REFERENCES**

[Beaudoin-Lafon, 2000] Michel Beaudoin-Lafon. Instrumental interaction: An interaction model for designing Post-WIMP user interfaces. In *Proceedings of CHI 2000*, pages 446-453, The Hague, The Netherlands, April 2000. Association for Computing Machinery.

[Clark and Marshall, 1981] Herbert Clark and Catherine Marshall. Definite reference and mutual knowledge. In A. K. Joshi, B. L. Webber, I. A. Sag (Eds.), *Elements of discourse understanding*, pages 10-63. Cambridge University Press New York, 1981.

[Clark and Wilkes-Gibbs, 1986]. Herbert Clark and Deborah Wilkes-Gibbs. Referring as a collaborative process. *Cognition*. 22:1-39, 1986.

[Lambert and Carberry, 1992]. Lynn Lambert and Sandra Carberry. Modeling negotiation subdialogues. In Proceedings of the 30<sup>th</sup> Annual Meeting of the Association for Computational Linguistics, pages 193-200, Newark, DE. Association for Computational Linguistics.

[Novick and Perez-Quñones, 1998]. David Novick and Perez-Quñones. Cooperating with computers: Abstracting from action to situated acts, *Proceedings of the Ninth European Conference on Cognitive Ergonomics (ECCE-9)*, pages 49-54. Limerick, IR, August, 1998. European Association for Cognitive Ergonomics.

[Shneiderman, 1983] Ben Shneiderman. Direct manipulation: A step beyond programming languages. *IEEE Computer*, 16(8):57-69, 1983.

[Suchman, 1987] Lucy Suchman. *Plans and situated actions*. Cambridge University Press, Cambridge, 1987.

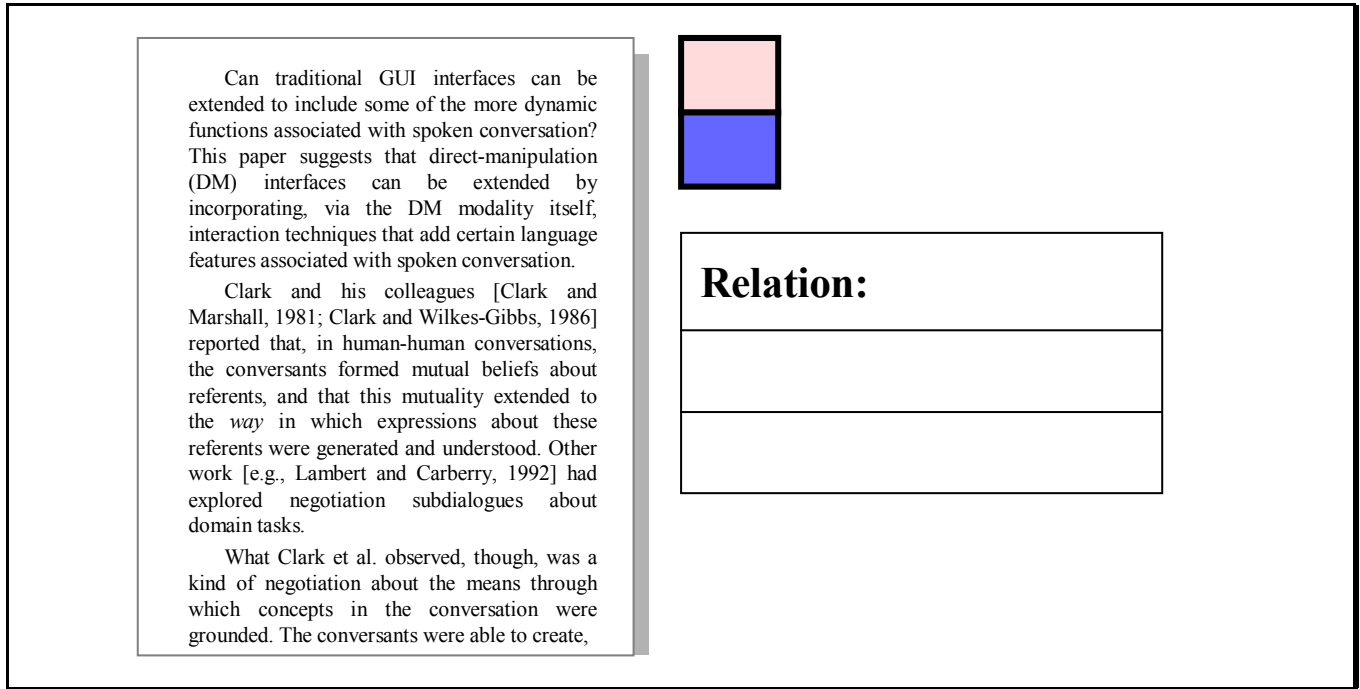


Figure 1. Text field with differentiated cursors

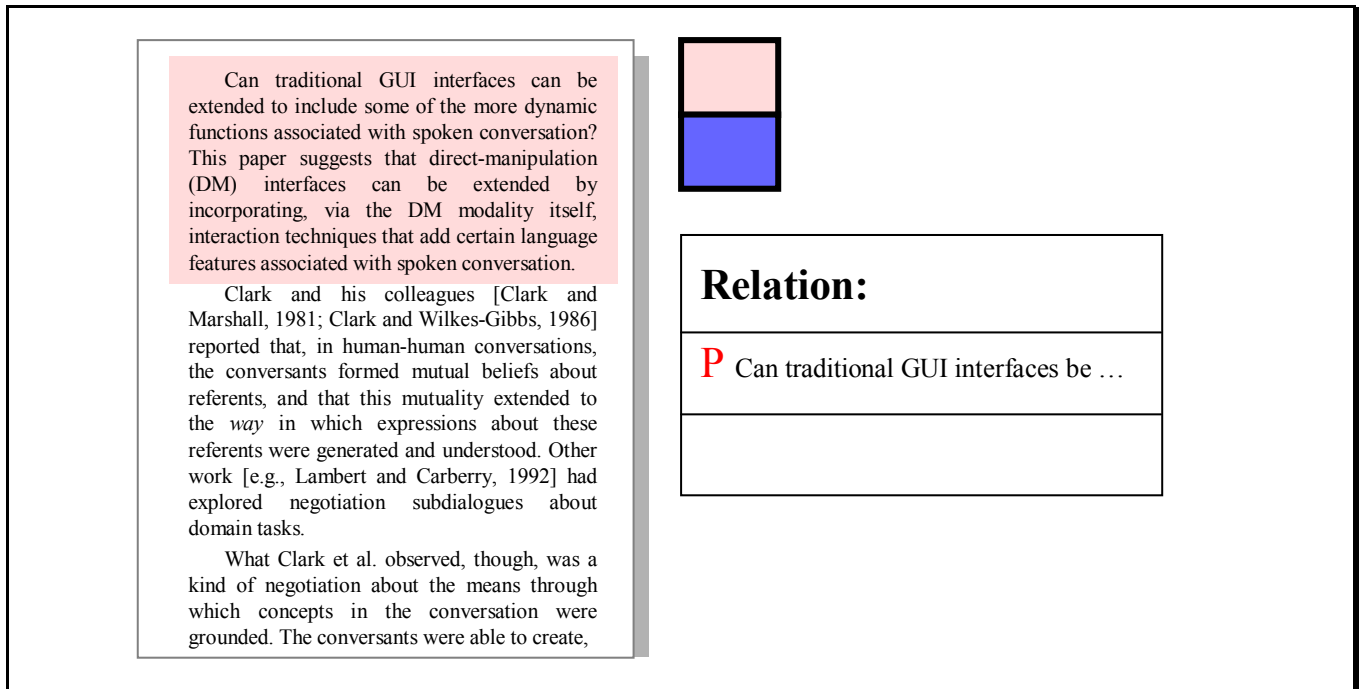


Figure 2. Text field with paragraph selected

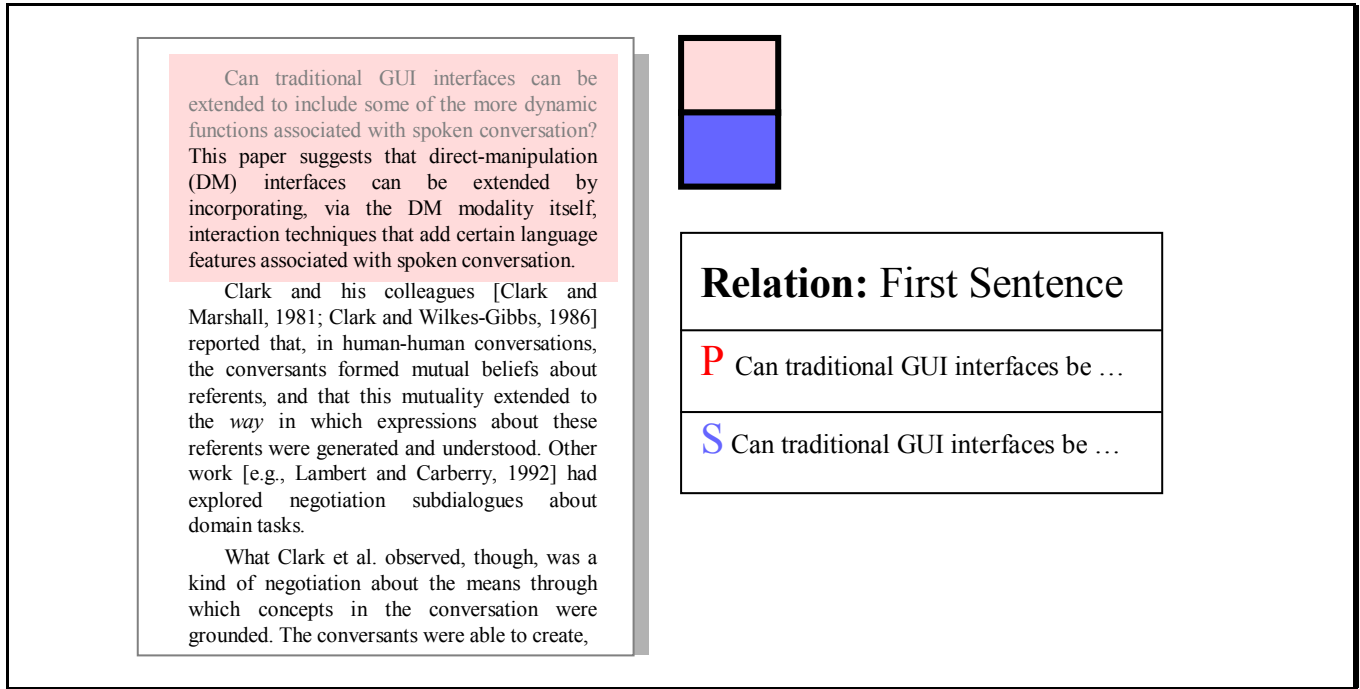


Figure 3. Text field with sentence selected within selected paragraph

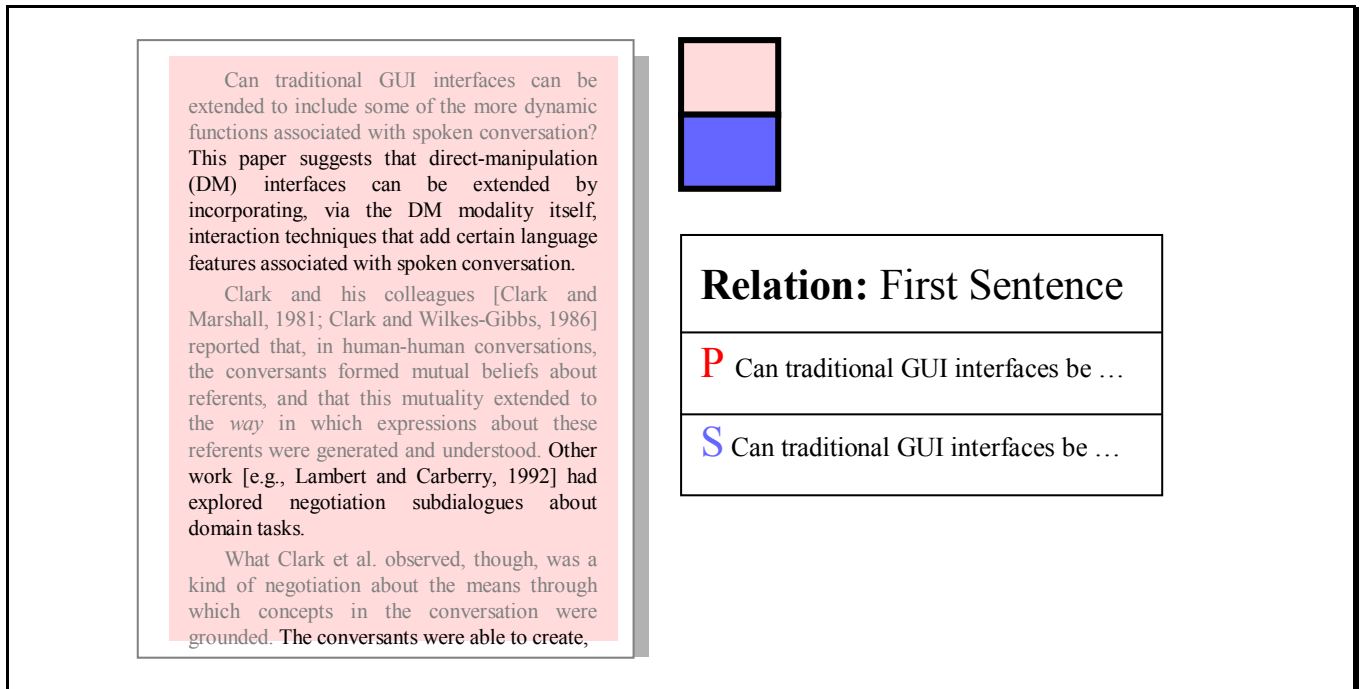


Figure 4. Text field with newly defined relation applied



This paper was published as: Novick, D., (2001). "Conversational" Dialogues in Direct-Manipulation Interfaces, *Proceedings of the 2nd IJCAI Workshop on Knowledge and Reasoning in Practical Dialogue Systems*, Seattle, WA, August 5, 2001