

Dialogue Cohesion Sharing and Adjusting in an Enhanced Multimodal Environment

Massimo Zancanaro, Oliviero Stock, Carlo Strapparava
IRST - Istituto per la Ricerca Scientifica e Tecnologica
38050 Povo TRENTO Italy
{zancana | stock | strappa}@irst.it

Abstract

Multimodality is a powerful concept for dealing with dialogue cohesion in a Human-Computer NL-centered system. This work is a little step in better exploiting the potential large bandwidth of communication provided by this situation. Light cognitive load feedback and direct manipulation are proposed so that user and system can cooperate in mutually conveying the cohesion of the on-going dialogue. Main points are: (i) use of an appropriate dialogue structure to constrain inference in the anaphor resolution process; (ii) use of a graphical representation of the structure, to limit the problem of opacity; (iii) allow the possibility of direct manipulation on this representation, to avoid the necessity of operating linguistically at the meta-level.

The context of the work is within a NL-centered multimodal information access system, in which basic entities are pairs (most commonly question and answer). A dialogue model simple enough to be shown in an intuitive-reading way on the screen, and powerful and accurate enough is provided by a modified version of the centering model.

1. Introduction

Extending the bandwidth of communication between human and computer, as in the case of multimodal systems, can open new scenarios in dialogue processing. The focus of our attention here will not be on the integration of a deictic context and visual presentations of domain entities, (see for instance [Maybury to appear], and [Stock 1992]), but on a different point: synthetic feedback and direct manipulation so that user and system can cooperate in mutually conveying the cohesion of the on-going dialogue. In general, the perspective envisioned in multimodality is one in which the balance of responsibilities in dialogue can be altered from the natural one. For some cases, in this connection, the active role of the hearer should be enhanced [Stock 1992], but in the case presented here something is added to the speaker's part to make the hearer's part easier.

In natural dialogues, while the use of pronouns is very convenient for the speaker, the hearer has a much more difficult task in identifying the desired reference; and of course the more so for an artificial NLP system. The problem bounces to the original speaker if there is no confirmation that the hearer has found the intended reference. Often, a solution provided by question-answering systems

consists in a paraphrase capability. However, paraphrases are boring, disruptive and often pragmatically inaccurate expressions.

Misconceptions are very common indeed. In natural life situations, another natural language process of communication has to be started to agree on the suspended, failed process, if a communication failure has occurred. The two uses of the same NL medium are obvious for us, as there is a narrow bandwidth of communication available (yet in face-to-face communication we make use of our hands, for instance to refer to things said at different moments in a metaphorical time scale). In a NL human-computer system it is quite possible to avoid using natural language for this kind of metacommunication, provided there is a balance between syntheticity and accurateness of conveyed information. In a number of situations a fast turnaround confirmation, on the basis of some key coherence elements (resulting from some accurate underlying processing) to be conveyed in a simple graphical format, will do.

Let us take as an example the following dialogue:

Q: *Who is the secretary of the head of the department?*

A: *<the system tells the name>*

Q: *Can I have her telephone number?*

A: *Please do not disturb the secretaries*

The dialogue is successful if Q meant that the last question was related to the preceding answer in a certain way. The key information on this point may suffice here as a feedback to Q so that he knows the utterance is being correctly interpreted. If he is expecting an answer such as

A: *The telephone number of the boss is.....*

it is advantageous that the system agrees with him on the key point that the sentence is to be related with the first Q's utterance in a certain way.

An important point for our human-computer system is that feedback is provided with a modality that involves light cognitive load, compared with the one involved in a complex clarification sub-dialogue or even in a paraphrase interpretation process by the human participant.

A shared representation that makes possible a natural-seeming recovery from discourse misconceptions by means of a series of "intuitive actions" is also wanted. So in the second situation of the above example, a simple check of

alternatives is appropriate that provides at a glance the right basic coherence relation. If approved by the speaker, this gives rise to the completely different interpretation.

If the system really is absolutely not able to get at what the user meant, there is a higher level of burden shared among the participants: in case of, say, a wrong resolution of an anaphor by the system, the user can point to the intended referent to force the right interpretation, with the system only providing the syntactically and semantically relevant candidates.

Among the other complex aspects, metacommunicative expressions like "let us go back to my previous question..." and references of all kinds can be seen in a new light, in the work presented here.

An idea along the same direction is presented in [Cohen 1991] and has been integrated in the Shoptalk system [Cohen *et al.* 1989]. Shoptalk was probably the first system in which multimodality was used to resolve meta-level problems: anaphor is transformed into a problem of pointing with the use of windows for managing different focus spaces. Here, the difference between communication and metacommunication is explicit. At the level of metacommunication, Cohen's principle of synergicity [Cohen 1991]: "to use the strengths of one modality to overcome weaknesses of another" is applied to good advantage. Natural use of anaphors is not substituted, nor does the user have all the burden of explicitly relating things. The system understands, feeds back in a cognitively light manner and communication may be adjusted with little effort by the user.

The ideas presented here have been exploited for the purpose of a dialogue module for the "concierge" (an automatic information access system for a visitor concerned with the organization and scientific production of IRST) of the MAIA system, that includes also a porter-robot and an electronic librarian. Figure1 shows how the dialogue module is integrated in the system.

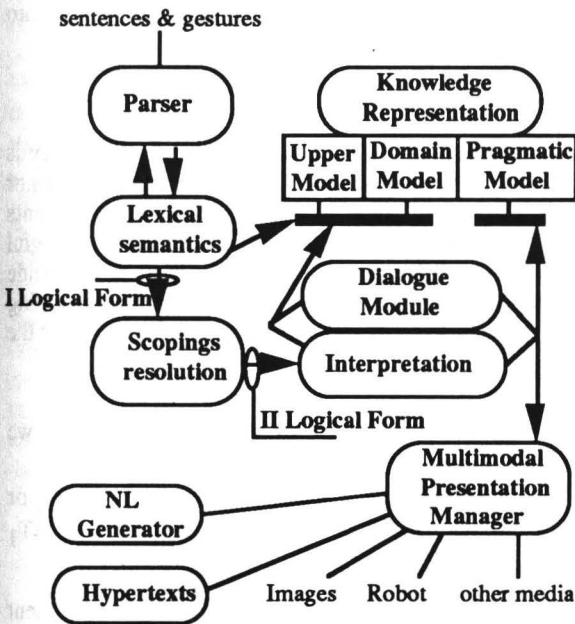


Figure 1

The main features of the design of this module are:

- (i) use of an appropriate dialogue structure to constrain inference in the anaphor resolution process;
- (ii) use of a graphical representation of the structure, to limit the problem of opacity;
- (iii) allow the possibility of direct manipulation on this representation, to avoid the necessity of operating linguistically at the meta-level.

To do this, a dialogue model simple enough to be shown in an intuitive-reading way on the screen, and powerful and accurate enough was needed. Here the centering model [Joshi *et al.* 1979], [Joshi *et al.* 1981], [Grosz *et al.* 1983], [Brennan *et al.* 1987], [Kameyama 1986] was revisited and modified for adaptation to a Q/A-system.

The next section briefly explains the basis of the modified centering model. Anaphor resolution will then be discussed; in the rest of the paper multimodality will be introduced.

2. The modified centering model

The centering model, first proposed by [Joshi *et al.* 1979], is based on the assumption that "in a natural language sentence in discourse [...] one or more entities [...] are singled out" [Joshi *et al.* 1981]. These entities, called *centers* are used to formally define the notion of discourse and to control the inference in the process of natural language understanding. This model was then refined by various authors (among the others [Grosz *et al.* 1983], [Brennan *et al.* 1987], [Kameyama 1986]) and used to resolve the problem of anaphor resolution. [Brennan *et al.* 1987], quoting a preliminary draft of Grosz, Joshi and Weinstein, proposes a set of typology of transitions to one utterance to the next to model the discourse structure and to resolve anaphoric expressions. The basic idea of this model is that to every utterance U_n a list of *forward looking centers*, $Cf(U_n)$, made up of all entities realized or directly realized ([Grosz *et al.* 1983]) in the utterance is associated. Ranking of elements in this list is crucial. In [Brennan *et al.* 1987] obliqueness of grammatical relations of the subcategorized functions of the main verb is proposed: subject, object, object2, then the other subcategorized functions and finally the adjuncts. Actually, for freer word order languages than English, like Italian, the position (in the surface structure of the sentence) of the subject with respect to the verb, or other syntactically marked constructions (topicalization, dislocation,...) must be taken into account.

The first element of the $Cf(U_n)$ is called *preferred center*, $Cp(U_n)$. It is the element for which is largest the likelihood of being the primary focus of the subsequent discourse.

U_n is "about" only one entity at a time, the *backward-looking center*, $Cb(U_n)$. This is always the confirmation of an already mentioned entity.

$Cb(U_n)$ is used to link U_n to the preceding discourse, while $Cf(U_n)$ provides a set of entities to which the succeeding discourse may be linked [Grosz *et al.* 1983].

The model of centering was developed in a framework of discourse analysis. For the case of dialogues and especially Query/Answer dialogues these notions seem not to be directly usable. The (implicit) assumptions of the centering

model are that utterances are mainly declarative and the purpose of the discourse is to explain facts. This is not true in a Q/A system in which half of the utterances are questions or commands and the purpose of the communicative scenario is to get some information about one or more subjects.

The version modified for use in a Q/A system, with the purposes outlined in the introduction, follows.

The minimal unit is considered not the utterance but the turn (i.e. the adjacency pairs Question/Answer and Request/Response). Centers are now properties of turns, not of utterances.

As before, $Cb(T_n)$ is used to link the turn T_n with a preceding one, and $Cf(T_n)$ to link it to the following ones. Among the forward-looking centers we choose one to be the $Cp(T_n)$. It is "the most significant" NP of the turn (i.e. the item realized by it), usually the wh-focus for wh-questions and the "thing about which we ask" in the other cases (cfr. paragraph 2.1).

We can informally think of a segment as a coherent chain of turns. In a Q/A dialogue we can imagine that a question is coherent with the preceding dialogue when it is a follow-up of a preceding turn. This is not the only case of coherence, of course, but it is a good starting point.

There are at least two ways in which a question may be a follow-up of a turn, either it is about the same subject of the question of that turn, or it is about the same subject of its answer. We call α -coherence the former and β -coherence the latter (α -coh and β -coh respectively, for short). The ambiguity found in the example in the introduction is precisely between these two coherence relations.

Formally:

If T_1 and T_2 are turns:

T_2 is α -coh with T_1 ($T_2\alpha T_1$) iff $Cb(T_2)=Cb(T_1)$

T_2 is β -coh with T_1 ($T_2\beta T_1$) iff $Cb(T_2)=Cp(T_1)$

T_2 is coherent with T_1 iff $T_2\alpha T_1$ or $T_2\beta T_1$

Now, a *segment* can be formally defined as a totally ordered set of turns among which some coh-relation holds.

(i.e. turns are totally ordered regarding the moment of their occurrence)

2.1. How to determine the centers of a turn

It is well established that in order to determine the centers syntactic considerations only are not enough [Grosz *et al.* 1983], but different criteria have not been agreed upon. In our case, with a specific applied situation, it has been possible to find reasonable strategies.

All items that are referred to by an NP in the sentence uttered by the user are used as centers. Some heuristics are then used to pick out the backward and the preferred center. These heuristics are very complex, and here the way they work is only outlined.

To determine the preferred center, it seems to be a good heuristics to consider the variable of the outmost quantifier in the logical form of the user's sentence. For example in the sentence

Who manages IRST?

(who ?x (person ?x) (manage ?x irst))

we choose the person identified by the variable ?x (centers are properties of turns, not of sentences, thus when the turn is closed the variable will be bound to an item in the knowledge base). This heuristics corresponds to choose the wh-focus. A problem with this strategy arises in the case of non-quantified sentences, for example

Show me the office of Stock?

(show (ref-term ?x (and (office ?x)(is-property-of ?x stock))))

In this case we suppose to be "stock" the Cb and the particular office the Cp.

Also for quantified formulas there are some problems, in particular with sentences with verb "to be" used in sense of IS-A and with the quantifier "where".

As an example for the first case, consider:

What is the organization of IRST?

(what ?x (thing ?x)

(isa ?x (ref-term ?y (and (organigram ?y)

(organigram-of ?x irst))))

?x and ?y are bound to the same entity but it is better to choose ?y as Cp because it has a stronger restriction (organigram vs thing). For "where", the problem lies in the semantics of this quantifier. For example

Where is the director's office?

(which ?x (place ?x)

(has-place (ref-term ?y (and (office ?y)

(is-place-of-person ?y

(ref-term ?z (director ?z))))

?x))

here, the wh-focus is the place, but it is not linguistically realized; the heuristics is to choose the term including the place, here the office described by the variable ?y.

As far as the backward center is concerned, there are two cases, whether the current turn is at the beginning of the segment or not. In the first case, the only candidates are constants and definite NP used in Russellian sense (iota-terms), otherwise pronouns, anaphoric demonstratives, definite NP and deictic demonstratives in this order, are sought. Actually, many peculiar cases are taken into consideration.

3. Anaphor resolution

One of the purposes of a discourse (or dialogue) structure is to limit the inference required to the process of anaphor resolution. The key idea is to consider the referents introduced in the preceding dialogue that could be meaningful at a certain time, only on the basis of their position inside the segment. The set of these referents is used as a starting point for the application of other criteria in order to select the correct ones.

Given two turns T_1 and T_2 belonging to a segment S, we say that

T_1 is accessible from T_2 iff T_2 is coherent with T_1 or exists a chain of turns $T_{k_1} \dots T_{k_n}$ where $T_{k_1}=T_2$, $T_{k_n}=T_1$ and for every i T_{k_i} is coherent with $T_{k_{i+1}}$ ($k_i > k_{i+1}$)

As an example, consider the tree in Figure 2 for segment $S=\langle T_1, T_2, T_3, T_4 \rangle$.

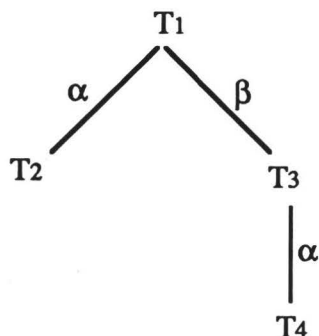


Figure 2

T_1 is accessible from all the other turns, but, for example T_2 is not accessible from T_4 .

If T is a turn belonging to a segment S :

The *Strong Referents Set* of T (T_{SRS}) is the set:

$\cup \{Cb(M), Cp(M)\}, \forall M$ such that M is a turn accessible from T

The *Weak Referents Set* of T (T_{WRS}) is the set:

$\cup C(M), \forall M$ such that M is a turn accessible from T (where $C(M)$ is the set of all centers of T)

Clearly, it is always true that $T_{SRS} \subseteq T_{WRS}$.

WRS and SRS are used as focus spaces in the process of anaphor resolution. Following the principle of factor cumulation¹, heuristics operating on these focus spaces are divided into two classes, Propose Filters (PFs) and Cut Filters (CFs).

PFs are applied for each type of anaphor (personal pronouns, definite NPs, demonstrative adjectives and pronouns). Using either Referents Set, depending on the kind of expression, the set of all possible candidates is proposed for each anaphor. Then, the set of all possible assignments of anaphors in the sentence is produced and the CFs are applied, cutting inconsistent assignments.

The reason for CFs to operate on the set of tuples is that the possibility for an entity to be the referent of an anaphor may depend also on the assignment made to other anaphors of the sentence, as seen in:

She gave her the permission to go home

Supposed that the results of the PFs are

she={boss, secretary}

her={boss, secretary}

there are four possible assignments

$\langle she, her \rangle = \{ \langle boss, secretary \rangle, \langle boss, boss \rangle, \langle secretary, secretary \rangle, \langle secretary, boss \rangle \}$

¹ "various clues from a given text work together in supporting a certain hypotheses about a possible antecedent for an anaphor and finally the correct hypotheses, indicating the intended antecedent, outweighs the other ones" [Carter 1987]

but, clearly, only the first is acceptable. The last is refuted because of world knowledge, the other two because of a well known syntactic constraint.

4. Direct manipulation of the structure

Using such a simple model of dialogue structure, it is possible to build what stated in the introduction as points (ii) and (iii).

It is self-evident, by the formation rules, that the structure is a graph with labelled arcs: nodes are turns, an arc between two nodes signals a coh-relation between the corresponding turns, the arc label indicates the type of the relation (i.e. α -coh or β -coh).

Since turns are totally ordered with respect to the moment of their occurrence, the formation of circuits in the graph is impossible. From the rules it is clear that every node in the graph has one and only one parent, except the one corresponding to the first turn of the segment. Therefore the graph is a tree.

To improve readability on the screen, instead of having labels, arcs are graphically characterized in different ways to show the kind of relation they represent. Labels are used to show the centers that have established the relation (i.e. backward for α -coh and preferred for β -coh).

Figure 3 displays a snapshot of a dialogue with MAIA. At the bottom of the screen the control panel is seen. On the left, there are five buttons to control the windows (for parser, answer, structure and turns tracker). These windows are controlled both by the system and by the user. A button to impose free interpretation ("interpretazione libera") and another one to impose the opening of a new segment ("nuovo segmento") are visible. A final button allows the system to provide more alternative *structural* interpretations ("assegnazioni"); in the snapshot, it is barred to signal that there aren't any more. Cb and Cp of the current turn are also shown. Here is a translation of the utterances on the left of the screen together with their Cb's and Cp's:

Q: *Is there a group that took part in MAIA ?*

A: *The Human Factors group.*

[Cb=(MAIA, maia) ; Cp=(group, human-factors g.)]

Q: *Who manages it ?*

A: *Slack.*

[Cb=(it, human-factors g.) ; Cp=(person, Slack)]

Q: *Which group, managed by Stringa, took part in the project ?*

A: *The Vision Group.*

[Cb=(project, maia) ; Cp=(group, vision g.)]

Both the second and the third turns are follow-up questions of the first one. The second turn established a β -coh while the third one an α -coh.

A series of snapshots displayed in the appendix may give an idea of the use of the system.

As far as possibilities of direct manipulation of the structure are concerned, the following features have been implemented:

- Closure of a segment.

The next utterance will be interpreted as if it were the first one of the dialogue.

- Visualization of an old segment.

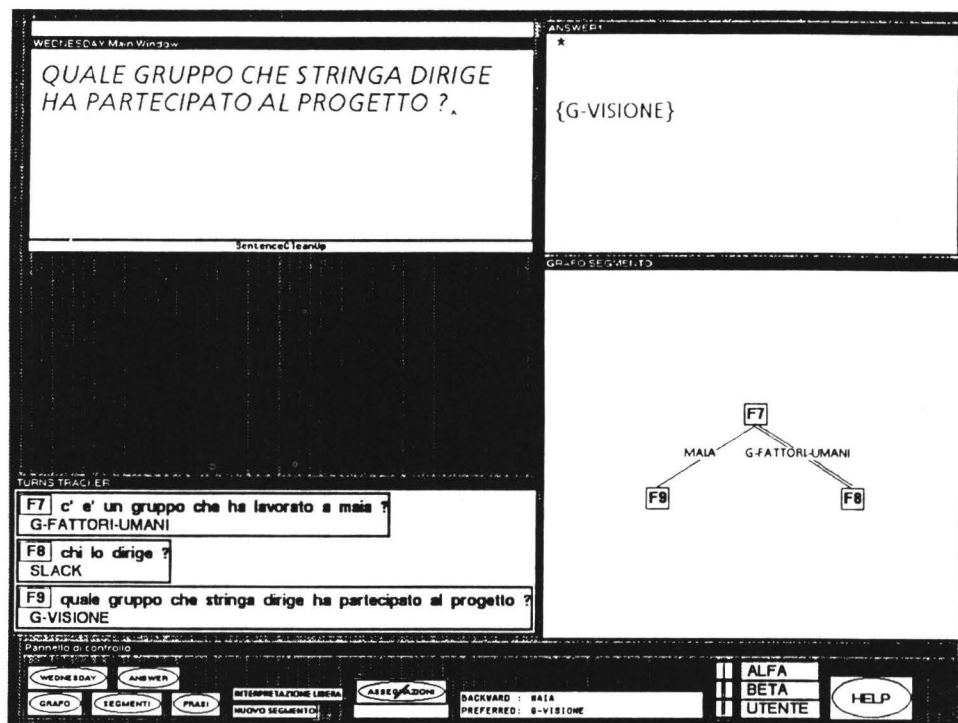


Figure 3

The system shows the tree and the corresponding linguistic realization of the segment (i.e. the questions/commands by the user and the corresponding answers by the system).

- Swap of the current segment with an older one. The older segment becomes the current one, thus the next utterance will be interpreted as if it were uttered there.
- Visualization of the point at which the user starts to travel through an hypertext.

An icon is automatically set every time the user chooses to follow an hypertextual link. This signals a momentary interruption of the flow of the dialogue.

- Check of alternative interpretations of the utterance, and choice of one of them.

If, at the end of the process of anaphor resolution, there are more than one assignment, the user has an immediate perception of the system's interpretations, by overlooking the simple structure. If this is in contrast to what he meant, he explores different alternatives, and may give the impulse to a complete interpretation with the new bindings.

- Build a new interpretation by imposing anaphoric resolution.

If none of the assignments proposed by the system is the one intended by the user, he may choose a new interpretation imposing "by hand" the resolution of anaphors. The resolution is guided by the system, highlighting the items that on the basis of selectional restrictions could be appropriate referents.

5. Taking into account the multimodal context

We have studied other "linguistic" coherence relations beside

those introduced above (for example those induced by elliptical constructions, renominalizations and so forth) but will not talk about this at any length here.

A related problem is the role of a deictic context in establishing coherence in a dialogue. A sequence of turns may be coherent even without any form of linguistic cohesion, just because utterances "talk" about the same situation. In this paragraph we explain how the modified centering model has been expanded to cover some of these cases.

In MAIA the deictic context is always established as an answer to a question. We mark these turns as *context creator's* (CC). Whether a turn is a context creator is a matter of choices depending upon pragmatics; anyhow at a given time there can be only one active context creator turn in the dialogue. Our claim is that every reference to the context could be dealt with, for coherence matters, as a reference to the current context creator turn.

For CC turns all the entities referenceable in the context they establish are added to their centers. This list is called Cd (for deictic centers) and a new relation, named δ -coh is defined.

Let T_m a CC turn, and T_n a non-CC turn following T_m in the dialogue:

$$T_n \delta\text{-coh } T_m \text{ iff } Cb(T_n) \in Cd(T_m) \vee Cp(T_n) \in Cd(T_m)$$

A natural refinement to the above definition is to specialize it into two relations, depending on which of the centers of T_n , either Cb or Cp, establishes the relation.

$$T_n \delta_\alpha\text{-coh } T_m \text{ iff } Cb(T_n) \in Cd(T_m)$$

$$T_n \delta_\beta\text{-coh } T_m \text{ iff } Cp(T_n) \in Cd(T_m)$$

With respect to the graphical usage of the structure as a non-linguistic feedback, it may be appropriate to "collapse" some relations in order to maintain the property of intuitive-reading and light cognitive load. For example, δ_{α} -coh could be shown graphically just as α -coh and δ_{β} -coh as β -coh. We find that a better solution is to impose the same graphical form but different colours to those pairs.

As far as direct manipulation of the structure is concerned, other possibilities could be added to the above, exploiting the new coh-relations. For example, even if at a given time only one turn is the current CC, two CC turns can well occur in the same segment (the last being the active one). Returning to the preceding context (and thus making active the former CC turn) is a meta-level operation, very complex to perform in NL. Using direct manipulation, instead, it can be sufficient to point to the corresponding CC turn on the representation of the dialogue structure.

6. Further research

Many other territories are to be explored in this area. Among these, we are particularly interested in the evolution of the notion of coherence in a multimodal environment for information access in which the user shifts between goal-oriented NL dialogue and hypertextual navigation (see also [Stock 1991]). As noted by [Waterworth et al 1991] browsing is opposed to querying along the dimension of target orientation. These could be seen as two different modalities to which the principle of synergistic use applies. In this context our effort may also lead to better overall cooperation between system and user in minimizing the exploratory phase.

From a behavioural point of view experiments need to be performed to evaluate both the use of linguistic resources in the user and the overall user's task execution when using a multi-level synergistic interface [see also Stock 1992] as described here. This work is a modest step in that direction. It aims to exploit better the large potential that computers offer for NL communication. For certain aspects we think this direction may result in something different from what we are used to in human-human communication, both face-to-face or at a distance (telex, telephone, text) [Cohen 1984], but still well tuned to our cognitive and linguistic capacity.

Acknowledgements

We would like to thank all the members of the Natural Language group at Irst for many useful discussions.

References

[Brennan et al. 1987] Susan E. Brennan, Marilyn W. Friedman, Carl J. Pollard, "A Centering Approach to Pronouns",

Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics., Stanford, Cal., 1987.

[Carter 1987] David Carter, "Interpreting Anaphors in Natural Language Texts", *Ellis Horwood Limited*, 1987.

[Cohen 1984] Philip Cohen, "The Pragmatic of Referring and the Modality of Communication", *Computational Linguistics*, 10(4), 1984.

[Cohen 1991] Philip Cohen, "The Role of Natural Language in a Multimodal Interface", *Technical Note 514*, SRI International, Menlo Park CA, November 1991.

[Cohen et al. 1989] P.R. Cohen, M. Dalrymple, D.B. Moran, F.C.N. Pereira, J.W. Sullivan, R.A. Gargan Jr, J.L. Schlossberg, S.W. Tyler, "Synergistic Use of Direct Manipulation and Natural Language", *CHI'89 Proceedings*, May 1989.

[Grosz et al. 1983] Barbara Grosz, Aravind K. Joshi, Scott Weinstein, "Providing a Unified Account of Definite Noun Phrases in Discourse", *Proceedings of the 21st Annual Meeting of the Association for Computational Linguistics*, 1983.

[Guindon 1985] Raymonde Guindon, "Anaphor Resolution: Short-Term Memory and Focusing", *Proceedings of the 23rd Annual Meeting of the Association for Computational Linguistics*, 1985.

[Joshi et al. 1979] Aravind K. Joshi, Steve Kuhn, "Centered Logic: the Role of Entity Centered Sentence Representation in Natural Language Inferencing", *Proceedings of International Joint Conference on Artificial Intelligence*, Tokio 1979.

[Joshi et al. 1981] Aravind K. Joshi, Scott Weinstein, "Control of Inference: Role of Some Aspects of Discourse Structure - Centering", *Proceedings of the International Joint Conference on Artificial Intelligence*, Vancouver 1981.

[Kameyama 1986] Megumi Kameyama, "A Property-Sharing Constraint in Centering", *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics*, 1986.

[Maybury to appear] Mark Maybury (ed.) "Intelligent Multimedia Interfaces", *AAAI Press*, Menlo Park, Ca.

[Passonneau 1991] Rebecca J. Passonneau, "Some Facts about Centers, Indexicals, and Demonstratives", *Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics*, 1991.

[Stock 1991] Oliviero Stock, "Natural Language and Exploration on Information Space: the ALFresco Interactive System", *Proceedings of International Joint Conference of Artificial Intelligence*, Sidney 1991.

[Stock 1992] Oliviero Stock, "A Third Modality of Natural Language?", *Proceedings of the 10th European Conference on Artificial Intelligence, invited talk*, Vienna 1992.

[Waterworth et al. 1991] John A. Waterworth, Mark H. Chignell, "A Model for Information Exploration", *HYPERMEDIA vol.3 no.1*, 1991.

Appendix

CHI LAVORA NEL GRUPPO ?

{BONVICINI}
{MICH}
{BELLUTTA}
{GIORDANI}
{ASTE}
{BARBO}
{LEONARDI}
{STRINGA}
{BOZZOLI}

F7
MAIA
G-FATTORI-UMANI
F9
G-VISIONE
F10

F7 c'è un gruppo che ha lavorato a mesi ?
G-FATTORI-UMANI
F8 chi lo dirige ?
SLACK
F9 quale gruppo che stringe dirige ha partecipato al progetto ?
G-VISIONE
F10 chi lavora nel gruppo ?
(BONVICINI MICH BELLUTTA)

ALFA
BETA
UTENTE
HELP

Here is a continuation of the dialogue in figure 3. It is shown how to use direct manipulation on the screen to correct a misinterpretation.

The user's question is "Who works in the group?" There are two groups in the preceding dialogue. The system chooses the latter one; a β -coh relation with its turn is established. The button for alternative interpretations is no longer barred.

The alternative interpretation takes the first group (g-fattori-umani) as the referent for the definite NP; a β -coh relation with the first turn is established. The graph is shown on a black background emphasizing that it is only a visualization (i.e. no impulse to a new complete interpretation yet).

CHI LAVORA NEL GRUPPO ?

{BONVICINI}
{MICH}
{BELLUTTA}
{GIORDANI}
{ASTE}
{BARBO}
{LEONARDI}
{STRINGA}
{BOZZOLI}

F7
G-FATTORI-UMANI
MAIA
G-FATTORI-UMANI
F10
F9
F8

F7 c'è un gruppo che ha lavorato a mesi ?
G-FATTORI-UMANI
F8 chi lo dirige ?
SLACK
F9 quale gruppo che stringe dirige ha partecipato al progetto ?
G-VISIONE
F10 chi lavora nel gruppo ?
???

ALFA
BETA
UTENTE
HELP

CHI LAVORA NEL GRUPPO ?

{SLACK}
{CONATI}

F7
G-FATTORI-UMANI
MAIA
G-FATTORI-UMANI
F10
F9
F8

F7 c'è un gruppo che ha lavorato a mesi ?
G-FATTORI-UMANI
F8 chi lo dirige ?
SLACK
F9 quale gruppo che stringe dirige ha partecipato al progetto ?
G-VISIONE
F10 chi lavora nel gruppo ?
(SLACK CONATI)

ALFA
BETA
UTENTE
HELP

After the user's confirmation, the system completes the interpretation and provides an answer.