WHAT DISCOURSE FEATURES AREN'T NEEDED IN ON-LINE DIALOGUE

Eleanor Wynn
Xerox Office Products Division
Palo Alto, California

It is very interesting as a social observer to track
the development of computer scientists involved in AI
and natural language-related research in theoretical
issues of mutual concern to computer science and the
social study of language use. The necessity of writing
programs that demonstrate the validity or invalidity of
conceptualizations and assumptions has caused computer
scientists to cover a lot of theoretical ground in a
very short time, or at least to arrive at a problem
area, and to see the problem fairly clearly, that is
very contemporary in social theory. There is in fact
a discrepancy between the level of sophistication
exhibited in locating the problem area (forced by the
specific constraints of programming work) and in the
theorizations concocted to solve the problem. Thus
we find computer scientists and students of language use
from several disciplines converging in their interest
in the mechanics and metaphysics of social interaction
and specifically its linguistic realization. Attempts
to write natural language programs delivered the reali-
zation that even so basic a feature as nominal reference
is no simple thing. In order to give an "understander"
the wherewithal to answer simple questions about a text,
one had to provide it with an organized world in which
assumptions are inferred, in which exchanges are treated
as part of a coherent and minimally redundant text, in
which things allow for certain actions and relations and
not others, and for which it is unclear how to store the
information about the world in such a way that it is
accessible for all its possible purposes and delivered
up in an appropriate way. Some of these were providable
and some weren't. Some AI workers have already moved
into the phenomenological perspective, just from con-
fronting these problems -- a long way to go from the
assumptions of mathematics, science, and engineering
that they originally brought to the task.

Others, in their attempts to deal with issues of repre-
sentation and motivation in discourse, have started
recreating segments of the history of social theory.
This is the history and perspective that students of
social interaction bring with them to the problem. They
arrive at the problem area either through a theoretical
evolutionary process in which they reject the previous
stage of theory, and interaction is a good demonstration
of the limitations of that theory, or because they are
simply intrigued by observing the wealth of social
action with which they can identify as members, that the
study of naturally-occuring discourse provides.

In social theory, the ethnomethodological perspective
arose as a response to the:

1) political implications
2) reifications
3) unexamined assumptions
4) narrow filter on observation

presented by structural-functionalist theory.

This theory:

1) limits and constructs observation fairly strictly
2) justifies the status quo (whatever exists serves
   a survival function)
3) posits a macro-organization (well-defined
   institutions and roles)
4) uses platonic idealizations of the social order
5) is normative
6) doesn't explain change very well

Difficulties in this theory were in part an artifact of a general positivist-scientistic orientation in which there was a motivation to treat the social world as a scientific object and hence to structure the description of it in such a way as to make the social world amenable to prediction, testing and control. The ethnomethodological or phenomenological perspective does not give up the scientific pretension but it does drop the engineering motivation. A world whose modus operandi (to avoid saying rules) or practices are constantly being created on the spot and which, though following along recognizable tracks, is in a constant state of invention and confirmation, lends itself far less to prediction. In fact it is clearly unpredictable.

Language itself provides an analogy, though it is partly the character of language that allows for the constant state of invention in the social world. Language changes constantly by means of several mechanisms, among which are phonological drift, usage requirements, metaphorization, and social emulation based on values and fashions. For theoretical purposes, one of the most valuable findings in Labov's landmark quantitative studies of phonological variation, was that social values drive the distribution of optional variants from one speech occasion to another according to the perceived formality of the occasion. In this manner, values -- what individuals at different social levels consider to be prestigious articulations, drive phonological change in general. Linguistic fashions themselves also change in response to what is currently used, and change with or against the majority according to the kind of identification desired to be made. They cannot be predicted in advance as such changes in value are typically discovered not planned. Very often changes in language use are derivative, based on a secondary or marginal meaning or usage, or discovered analogy or metaphor of some existing locution. Thus a dynamic of social contrasts and identifications, as well as social mobility and aspirations thereto, as well as socially situated invention, are deeply connected to linguistic issues, including language change and the concept of distribution rules, in an empirically observable and countable way. These and other social dynamics operate no less for more complex discourse phenomena, and account for large portions of observed discourse strategies.

Generally, when a sociolinguist, sociologist, or anthropologist looks at language use, what they attend to are the disclosed social practices. Being aware of, and focussing on social context, with a history of social theory or an historically developed set of concepts for social action in mind, alerts one to many attributes of the occasion for interaction: the possible social identities and relationships of the participants, the perceived outcomes and the social significance of meanings generated in the course of the interaction, as well as to structural and habitual features that reflect social requirements (viz. the "recognition" requirement as a prerequisite to interaction's taking place at all or in the particular form, as discussed by Schegloff).

The fact that a background of shared knowledge about the world is assumed emerges from an examination of what is explicitly stated and from the observation that what is explicit is in some way "incomplete", partial, not a full itemization of what is communicated and understood. It is also the case that to spell out all the assumptions would be unbearably time-consuming, redundant to the purpose, boring, and possibly an infinite regress; and this practice would moreover fail to accomplish all those conversational purposes which require negotiation, building up to a point of mutual orientation and accord, or the "use" of one person by another for a real or imaginary gain. (cf Simmel)

The messiness, potential ambiguity, implicitness, etc. of natural conversation serve many of the purposes that actors have, including the one of intimacy and mutuality by less and less explicit surface discourse. Herein lies an important distinction, one that is not well perceived by workers in AI. Purposes can be, and typically are discovered in the course of interaction rather than planned. Purposes are thus emergent from interaction rather than apriori organizing principles of it.

Attempting to code, catalogue, regulate, formalize, make explicit in advance those purposes is reminiscent of structuralist, positivist social theory. To this extent, computer scientists are recreating social theory, starting from the point that is most amenable to their hopes and needs, and so far lacking the dialectic that contextualizes other developments in social theory. Ontogeny has not yet fully recapitulated phylogeny. Extending the plans, goals, frames notion into the wider social world (wider than a story understander), constitutes a platonic idealization and the ensuing problem of locating those idealizations somewhere, as if there were large programs running in our heads (some of which need debugging), or as if there were some accessible pool of norms from which we draw each time we act. It posits that we act out these idealizations in our everyday behavior, that our behavior constitutes realized instances of this structure. This conflicts with a "process" notion of interaction, which careful discourse analysis reveals, whereby participants are continually trying out and signalling their participation in a mutual world, presumably because this is not from one instance to the next pre-given. The great revelation of discourse analysis in general, if I may be so sweeping, is the ability to observe the process of social action, whereby the social world is essentially built up anew for the purpose at hand, and interactants can be seen sorting out the agreed-on premises from those that need to be established between them.

There are two kinds of concerns here that bear upon on-line dialogue research. One is the notion of person, social identity, etc. The other is the notion of interaction as a reality testing mechanism that grounds the individual in a chosen point of view from among the many interpretations available to him for any given "event". Both of these notions differentiate the computer from a person as an interactant. Sorting out dialogue issues that embody these notions, narrows down the field of concerns that are relevant for building "robust" on-line dialogue systems.

All social systems, including non-human ones, display social differentiation. This is a central notion that the AI path of evolution does not bring to the study of discourse. On the contrary, discourse problems are treated as if there were a universality among potential interactants. This fits very nicely with a platonic perspective. Kling and Scacchi have referred to this as the rationalist perspective, and they cite claims made for simulation and modelling as their illustration of how exponents of this perspective fail to make even gross social distinctions:

"Neglecting the obiter dicta claim that modelling and simulation are 'applicable to essentially all problem-solving and decision-making,' presumably including ethical decisions, one is left with an odd account of the problem of modelling. Models are 'far from ubiquitous' and 'the trouble is' they are difficult and costly to develop and use. But the appropriateness of modelling is not linked by (rational perspectivists) to any discernible social setting or the interests of its participants. (Their) claims are not aimed at policy-making in particular. They could include simulations

for engineering design as well as for projecting the costs of new urban development. However, their comments typify the rational perspective when it is applied to information systems in policy-making; the presumption is that differences in social settings make no difference."

Work in socio-linguistics, on the other hand, has focussed on how speech varies by situation, by relationship, by purpose and by many other constraints that depend upon both a typification of the other from a complex set of loose attributes and the discovery of his unique behavior in the situation. The notion of a linguistic "repertoire" expresses people's demonstrated ability and propensity to adjust their speech at almost every analytic level, down to the phonology, to their perception of the situation and the audience. There are variations in people's skill at this, but all do it. To the extent that they don't do it, they risk being inappropriate and not getting rewards from interaction. (see F. Erickson for a study of the outcomes of interactive strategies in ethnically mixed interactions.)

The structuralist perspective again may be an appealing way for computer scientists to approach the problem of differentiation of persons, as it posits an essentially limited set of "roles" of fairly fixed attributes, and posits as well an ordered hierarchical arrangement of those roles. With this framework in mind it is relatively easier to imagine a computer as a viable participant in a social interaction, as it should be possible to construct an identifiable role for it. With this rather flat view of human social perception it is also possible to imagine a person requiring of a computer that it behave appropriately in a conversation, without regard for the fact that a computer can only satisfy a very limited set of purposes for that person in interaction. In fact people know perfectly well many of the things computers can't do for them or to them, things which other people can do and hence which need to be taken into account in dealing with other people. And they are able to differentiate for the purpose of interaction among infinitely many people, and states of mind or situation those people can be in.

The other feature of interaction between people, reality-testing, is less well understood than differentiation, which is a veritable solid ground of social understanding. However, it can be seen in interactions, even very simple task-oriented ones such as I described in my thesis, that people are also always accessing each other for a view of the world, for agreement, disagreement, and a framework for interpreting. Diffuse explanation mechanisms(Wynn, 1979) also exhibit the tendency of speaker to nail down the audience's perception of himself to the framework of interpretation desired by him, as an implicit acknowledgement of possible variance. What is often uncertain in an actor's "model" or projection, or understanding of the other participants or observers, is their view of the actor himself. To this end, he fills in and guides the interpretation with additional context any time he perceives an occasion for misinterpretation, sometimes to the point of logical absurdity (but practical appropriateness if not necessity).

Since a computer is not an actor in the social world, its interpretations, both of oneself and of "events" -- perceived social phenomena-- don't really count. A computer can provide facts about the world within a well-understood framework, but it cannot provide the kind of context that comes from being a participant in social life, nor a validation of another's perception, except to the extent that matters of "fact" or true-false distinctions allow this. And in these cases, the person supplies this validation himself from the information. This may be a moot point, but I maintain that the search for agreement, confirmation, etc., and the related search for common ground or reality are basic motives for interaction, along with confirmations of membership and solidarity etc., as described in the work of Schegloff and of much earlier writers like Malinowski and Simmel.

Rather than working from careful and detailed observations of the real world, excepting such innovators as Grosz and Robinson, many computer scientists exhibit a tendency to develop their "models" of interaction by conceptualizing from the perspective of the machine and its capabilities or possible capabilities. Discourse features may be selected for attention and speculation because they offer either a machine analog or a machine contrast. Thus we people are attributed information structures, search procedures and other constructs which are handy metaphors from the realm of computerdom; and it would be especially handy if we were in fact constructed according to these clean notions, so that our thinking and behavior could be modelled. (In all fairness, I know computers have "guys" running around inside them, "going" places, "looking for" stuff, trying out things, getting excited or upset, going nuts, giving up, etc.)

Working from the machine perspective can lead to some gross observational oversights, and the authors of the oversight I've picked as an example will hopefully indulge me. The implicit confirmationhypothesis (Hayes and Reddy) could never have been hypothesized by anyone who studies language behavior from a social perspective, as one of the oldest conversational observations around is the explicit confirmation observation. The phatic communion notion is over 30 years old, and is perhaps the first attention given to those features of interaction whichwere initially considered to carry little or no observable propositional content or information. Included in these behaviorsare those discourse "fillers" that signal to the speaker he is being received with no problem, that the listener is still paying attention (even more basic than confirming), and that the listener is a participant in the rhythm of the interaction even though he is producing little speech at the moment. The "rights" and "hehhehheh's" of the current natural conversation transcription conventions are absolutely pervasive and omnipresent. Nods, "hm's", gaze, prompt questions, frowns, smiles, exclamations of wonder, are all explicit confirmation devices constantly used in conversation, and occur especiallywhennew propositions or details essential to building a story are presented. Speakers are also often tentative and reformulate at any evidence of withheld confirmation, like a "blank stare" or a frown from the audience.

Therefore it is by no means ungraceful to explicitly confirm, and on the other hand, it takes very little to do so. But the point is this: even if the implicit confirmation hypothesis were true (and I pick it because it is an available example and very easy to reject-- other notions would do as well but require a more detailed attack), it would be no reason to exclude this feature from a computer dialogue nor to suppose that it would pose people any difficulty in handling a dialogue with a machine. The discourse supporting activities of natural conversation always address practical concerns. If a new concern should arisebecause of newconstraints-- e.g. that the interactant is a machine--these will be incorporated in the ongoing details of communication. For instance, when it is obvious someone is having difficulty speakin- and understanding English, we unhesitatingly drop all ellipsis and give full articulation of every sound, even though this produces great redundancy in the message for purposes of communicating with another native speaker, and is moreover extremely unhabitual.

In fact, the social role of the computer is perhaps most like that of a foreigner. We assume a foreign individual whose English is poor to have an ability to communicate, perhaps a rudimentary grammar and vocabulary of our language, and a set of customs, some of which overlap with ours. But we can't take the specifics of any of these things for granted. There is very little in the way of a background of practices or assumptions to work with. But here the analogy ends.

Presumably, we won't be going to on-line dialogue programs to chit-chat. The purposes will be fairly well-defined and circumscribed. People will interact with a computer:

1) because there is no person available
2) because there is limited social confront in accessing expert information from a computer, so it is available in a metaphorical sense
3) because the computer has specialized abilities and resources not found in a single individual
4) because it coordinates non- local information and
5) is maximally up-to-date -- changes in status and the news of this are concurrently available and
6) the outcome of one's own interaction with the system may be animmediately registered action, like reserving a space and hence making one less space available to subsequent users
7) because actual searching (as opposed to the metaphoric kind attributed to our minds by cognitive scientists) of a large database may be required and the computer is much better and faster at this than we are.

In other words, our reasons, certainly our most solid and fulfillable reasons, for consulting computersand engaging in discourse with them will beto find out things relating to a framework we already have. The computer needs to know a few things about us and especially our language, and especially needs to know how to ask usto clarify what we said, even to present menus of intentions for us to choose from as a response to something unexecutable by it. But more than anything, it needs to be able to make its structure of information clear to us. In this sense it will satisfy certain "person" properties -- we have working notions of at least the parameters and starting points for negotiation with people. Whereas with computers we have at best an entry strategy for an unfamiliar system, but very little to go on in common knowledge for assessing its informedness or even consistency.

So on-line dialogue should not be like person-to-person dialogue in many respects. For instance, being overly explicit with a person is an indication of a judgment we have made about their competence. This judgment is quite likely to be offensive if it's wrong. (Schegloff) This is not likely to be a problem with a computer from an experiential social action point of view. Who cares if the computer cannot perceive that we are competent members of some social category defined bya more or less common body of knowledge: We will have no problem in telling it what level to address in dealing with us, if it has any such levels of explicitness, nor in gearing our own remarks to the appropriate level once we find out what it can digest. On-line dialogue systems therefore have an ongoing task of representing the - selves, not the whole interactive world; and designers need not concern themselves so much with providing their systems with models of users, but rather providing users withclear models of the system they are interacting with. These are the major concerns, obviously.

I wish I could now deliver the part of the paper that would be of most interest: what a dialogue system should contain and how it can make available those contents in order to realize the purposes just stated. Instead I have addressed myself to what look like common fallacies that I see in attempting to incorporate natural language dialogue issues into computer dialogue issues without access to the social understandings embedded in social interaction research.