

Modelling devices and modelling speakers

T.J.M. Bench-Capon and A.M. McEnery

The roles played in an illocutionary act by models of the means of communication and the communicator are distinguished, and qualitative differences between the models appropriate in the two cases identified. Applied to human-computer interaction, this means that a user must have models of the computer both as a communications device and a communications medium, and of the system author as interlocutor.

Keywords: speech acts, illocutionary acts, user models, human-computer interaction

This short note is a response to the commentary (Barlow *et al.*, 1989) on our views that computer dialogues should be seen as mediated discourse (Bench-Capon and McEnery, 1989). One point needs clarification: the commentary stated that we restricted our discussion to nonintelligent, nonadaptive systems. While finding the use of 'intelligent' somewhat question-begging, we did intend to include systems of all kinds currently found, including expert systems, knowledge-based systems, hypertext systems, and anything else that might be termed 'AI'.

Even so, the commentary does take a rather more inclusive view of interaction than was the concern of the original paper. In that paper we were addressing only the question of how a user should go about the interpretation of messages from the computer system, and the selection of input to that system.

It may be helpful to explain our reason for this focus. It was essentially that the elements of the interaction in which we were interested were those concerned with communication. We should perhaps at this point make clear our views on the nature of communicative acts. Following Austin (1962) and Searle (1969), we believe that an act of communication involves the simultaneous performance of a number of distinguishable acts. There is the 'utterance act' which involves the making of certain sounds, in the case of speech, or certain marks, in the case of writing, or causing certain images to appear on a computer screen, as in programming. Second there is a 'propositional act' in that these

Department of Computer Science, University of Liverpool, PO Box 147, Liverpool L69 3BX, UK. Tel: (051) 794 3697. Email: sq35@uk.ac.liverpool.ibm

sounds, marks or images may have some conventional semantics assigned to them by virtue of being part of some language. Third there is the illocutionary act in that the performer of the utterance act will have some intention to convey a meaning to the recipients of the act. This meaning is the illocutionary force, which may be distinct from the propositional force, the literal interpretation of the utterance. To complete the picture, there is the 'perlocutionary act', by which the recipient of the utterance ascribes a meaning to the utterance, which may be termed the 'perlocutionary force'. A successful act of communication occurs when the illocutionary and perlocutionary forces coincide. An act of communication therefore requires both an originator of an utterance, whom we term the 'speaker' irrespective of the communication medium, and one or more recipients of the utterances (the 'hearer(s)'). It is part of the responsibility of the speaker to make his utterance in such a way as to best promote the coincidence of the illocutionary and perlocutionary forces. This will influence the utterance act both in the way it is performed, perhaps suggesting the use of block capitals for some handwritten information, and the utterance made, since a variety of utterances with different propositional forces could be used to perform a given illocutionary act. Again in the original paper we concentrated on the latter aspect, giving stress to the selection of utterance taking into account the intended hearer and his cultural and linguistic propensities.

Thus our focus was on the interaction as an act of communication, posing the questions as to who is the speaker and who is the hearer, and offering programmer and user as answers. This was because we were interested in this aspect of the interaction and not because we wished to deny that any interaction at all with the device took place: to draw an analogy with telephone conversations, the act of communication is between the two people at either end of the telephone line. The telephone is a necessary condition of their communication, and they must understand how to use the telephone effectively, and must interact with their devices, but the communicative interaction that is taking place remains between the conversationalists. The role of secondary producers is a side issue with respect to communications aspects; the enjoyment of a telephone conversation may be enhanced by the ergonomic design of the instrument, but it contributes little to the understanding of what is being said. So too with books; a reader needs to have a model of the device so he can find particular statements, and may relish the fine binding and pleasant typeface, but understanding the content, ascribing a perlocutionary force, depends most crucially on the reader's model of the author. When we meet the phrase 'Mr Elton actually violently making love to her' (Austen, 1816) in a novel, it cannot be understood in the absence of a model of the author: we must envisage an altogether less physical scene knowing that it was written by Jane Austen than we would if it had been written by a contemporary author such as Jackie Collins.

Again one must not, however, read us as denying that the model of the device has any role in the speaker's selection of utterance and the hearer's interpretation of that utterance. A wise speaker will avoid the use of a word

which has a near homophone which could lead to misunderstanding when talking but need not take this trouble when writing. The use of the telephone places well-known constraints on useful forms of utterance; the lack of visual clues from the speaker may dictate a quite different form of words from what could be used in face to face conversation. Selection of utterance is a complex matter and must be influenced both by the speaker's understanding of his audience, which we stressed, and the speaker's understanding of the medium, which we did not. Similarly, interpretation is a complex matter to be performed in the light both of the hearer's understanding of the speaker, and the hearer's understanding of the medium. We remain of the opinion that the disregard of the medium is, however, more by way of a hindrance to effective communication, whereas disregard of the partner in the communicative act is likely to have disastrous consequences. Thus the model of the medium may be seen as being of a secondary importance.

Where the act of communication is mediated by a device, a model of the device is required for two different reasons. First, and most obvious, is that a model is needed because otherwise the participants in the interaction simply could not use the device to communicate. It is vital to know where to speak into a telephone if one is to be heard. This aspect of the device model assumes a greater importance in the case of computers than books and telephones. That is because we are all familiar with telephones and books (of a variety of sorts: novels, text books, directories, etc.), and they are so standard in their behaviour that we can use them without needing to develop a fresh device model for each telephone or book encountered or, indeed, without thinking about the model to any great extent. Only when confronted with a particularly novel sort of telephone do we need to consider how to use it. Computers, in contrast, are neither so familiar, nor so standardised, and have a potentially far greater range of functionality. Therefore when confronted with a new computer system it may well take us some time and experimentation before we develop our device model, and understand how to use the system effectively. Systems with WIMP interfaces provided a good example of this: when they were first introduced even experienced computer users took some time to become accustomed to the idea of opening icons with double clicks, dragging an icon around the screen by holding down the mouse button, and deleting files by dragging their icon to the mysterious picture at the bottom of the screen. Their existing models of the computer as a device were useless as a guide to the behaviour of the device with this extended functionality, and so needed extension to accommodate this extra functionality.

The second role of the device model is that it provides an understanding of the constraints operating on communication via the device. If communication is to be effective, these constraints must be understood by both participants in the communicative act, so they can work within them. Both of these roles together, however, are not sufficient to enable effective communication: ultimately successful understanding still requires that the hearer understands the speaker sufficiently to ascribe a meaning to the utterance.

One concrete example may help here. Suppose a user is stuck while interacting with a system, knowing what he or she wishes to do, but unsure how to do it. The user must be aware, from knowledge of the constraints imposed by the medium, that it is necessary actively to seek help; while a puzzled look might elicit help in a face to face situation, it is of no avail in this kind of mediated situation. Next, also from the device model, the user must be aware of how to solicit help on the topic causing confusion. Assuming that the device model is sufficient to enable the user to reach this point, it will now be possible for the user to secure some help message. But this message needs to be interpreted, and here the user must employ some model of the originator of the message, namely the programmer.

Users thus require models both of the computer, as a device and as a medium of communication, and of the system author, as speaker. The programmer requires similar models of the computer, and a model of the user as hearer. It is, however, vital to recognise that these models are qualitatively quite distinct. The models of the device are not only playing different roles from the model of the speaker, but are of an entirely different nature. The model of the system as a device is essentially there to answer questions of how to achieve certain effects, and comprises of a number of causal hypotheses of the form 'if I do this, then this will happen'. These hypotheses are meant to be consistent, so that the same cause is expected to result in the same effect. The model of the system as means of communication is quite different: this comprises a number of constraints on the interaction resulting from the features of the medium as compared with other potential communications media. Thus this kind of model inhibits forms of utterance dependent on visual clues from use over a telephone line, and forms of utterance depending on tone of voice from writing, and so on. Obedience to the constraints is, however, a matter of judgement: irony often depends on visual clues for its detection, but is not wholly impossible over the telephone, provided one is well known to the person with whom one is speaking. The mechanism of cause and effect is not paramount here. Models of speakers and hearers are different yet again; while there are strong causal elements here in that it is used to answer questions of the form 'how can I best explain this to my hearer?' and 'what made the speaker use that expression?', we have no expectation of consistency such as we have with devices. Use of the speaker/hearer models is an art, while use of the device models can be a science. This may just be because of the complexity and mutability of speakers and hearers as compared with devices, but it needs to be recognised by the users of such models.

The remark towards the end of the commentary (Barlow *et al.* 1989) about the probable cognitive impossibility of a user modelling the producer of the system probably stems from a confusion of these two kinds of model. A definitive causal model, in the sense of knowing which buttons to press to evoke a particular response from a person, may well be impossible. This cannot, however, be the case for a model of the sort appropriate to a speaker, since were it impossible, it would be impossible to understand any communication. Such a model is a necessary prerequisite of any communicative function, as we

attempted to argue in our original paper, and as is supported in Grice (1982), Diaper (1988) and elsewhere. Such a model must take into account any effect of the medium on the message, but while the characteristics of the device being used may be of some significance, the illocution will still depend for its success on the models held by the communicants of one another. All users of a language are more or less good at forming such models, and form such models whenever they participate in an interaction with another person. Nor is there any reason why the 'best professional psychologists' should be any better than anyone else at forming these kinds of models. The skill required here is not an understanding of the mind in a causal sense, but a social skill everyone develops and hones from infancy.

To conclude, our original paper focused on one particular aspect of human-computer interaction, namely the interaction as a communicative act. We did so because we felt that this aspect had been neglected or misunderstood: had we been trying to be uncontroversial we might have better titled our paper 'People communicate through computers, not with them'. Of course it is necessary to consider the role of the device in the interaction, and the constraints that it imposes on the communication. We do, however, believe that it is fruitful to see the interaction fundamentally as a communication between programmer and user, since this enables the distinctions made in philosophy alluded to above to be exploited within the discipline of human-computer interaction. It is worth considering as separate issues how to aid users in constructing device models so they can know how to use the medium to communicate, how the utterance should be formed (particularly as the computer offers such a range of options such as text, pictures, and sound), what constraints the mediation of the computer places on the forms of utterance, and how the user can be aided in constructing a model of the speaker. Our view is that these distinctions are currently blurred, to the disservice of all concerned. Finally it would assist the designers of systems if they recognised themselves as speaking to their users through the computer rather than as building a system that would speak to users on its own account. We trust that this note helps make clear the distinctions that underlie our views, and corrects any overstatement in our original paper.

References

- Austen, J. (1816) *Emma* Thomas Nelson, London, 116
- Austin, J.L. (1962) *How to do things with words* Oxford
- Barlow, J., Rada, R. and Diaper, D. (1989) 'Interacting WITH computers' *Interacting with Computers* 1, 1, 39-42
- Bench-Capon, T.J.M. and McEnery, A.M., (1989) 'People interact through computers not with them' *Interacting with Computers* 1, 1, 31-38
- Diaper, D. (1988) 'Natural language communication with computers: theory, needs and practice' in Duffin, P. (ed.) *KBS in Government 88* Blenheim Online, Pinner, UK, 19-44
- Grice, H.P. (1982) 'Meaning revisited' in Smith, N. (ed.) *Mutual knowledge* Academic Press, London, UK
- Searle, J.R. (1969) *Speech acts* Cambridge University Press, Cambridge, UK