

Talking with pictures: exploring the possibilities of iconic communication

Colin Beardon*, Claire Dormann*, Stuart Mealing** and Masoud Yazdani**

*Faculty of Art, Design and Humanities, University of Brighton, Brighton BN2 2YJ

**Department of Computer Science, University of Exeter, Exeter EX4 QH

Abstract

As multimedia computing becomes the order of the day, so there is a greater need to understand and to come to terms with the problems of visual presentation. This paper deals with iconic languages as a means of communicating ideas and concepts without words. Two example systems, developed respectively at the universities of Exeter and Brighton, are described. Both embody basic principles of the iconic communication which, though not unique to learning technology, is forming an increasingly important part of user-interfaces, including those in the area computer-assisted learning.

Introduction

Koji Kobayashi looks forward to a time when the telephone system will automatically translate between the languages of users in real time (Kabayashi, 1986). It will then be possible to speak in your native tongue to a Frenchman, a Japanese and a Russian, and for each of them to hear your words spoken in their own language as you speak. In *The Hitch-Hiker's Guide to the Galaxy* (Adams, 1979) it becomes possible to understand communication in any language by plugging a Babel fish into your ear – a fish with the natural ability to receive speech in any language and translate it into the language of the person wearing it. While these visions are certainly futuristic, even fantastic, the need for people knowing different languages to communicate quickly and easily is rapidly increasing at the same time as the technology for transmitting and receiving information is spreading and becoming more powerful.

Attempts to create international languages have not been very successful, partly because of the need for a significant number of people to know them, and partly because they have to be learned like any other new language. Understanding is, however, possible across language

barriers in at least two ways. First, there are a number of internationally recognized signs, symbols and icons which one can find on the roadside and in airports, railway stations and similar places. They are not only used directly to denote the place where one might find trains or change money, but may also contain a number of 'meta-icons' such as the red diagonal and the arrow. Secondly, it is possible to make oneself understood, at least at some basic level, by means of gestures and mime. People who find themselves in a country where they know little or nothing of the local language can often communicate simple ideas by pointing or indicating their intentions through more performative signs, which may border on acting.

Iconic communication is the attempt to build cross-language communication systems¹ that completely avoid the use of words and rely solely on pictorial symbols. The system which manipulates these symbols should be as simple to use as possible, though it is clear that it has to be learned in some way. There are various forms of learning which can apply: users can receive some form of instruction (either by a tutor, or by some non-linguistic tutorial system); the elements of the system can contain an explanation of their meaning ('self-explaining icons'); the system can adopt a powerful metaphor; or the user can experiment and learn by trial and error.

Icons

It is important to say first of all what we mean by an icon. Figure 1 contains various ways of referring to 'men', and they vary from the arbitrary (Figure 1a) to the pictorially descriptive (Figure 1d). Icons tend to exist between these extremes (Figures 1b and 1c).

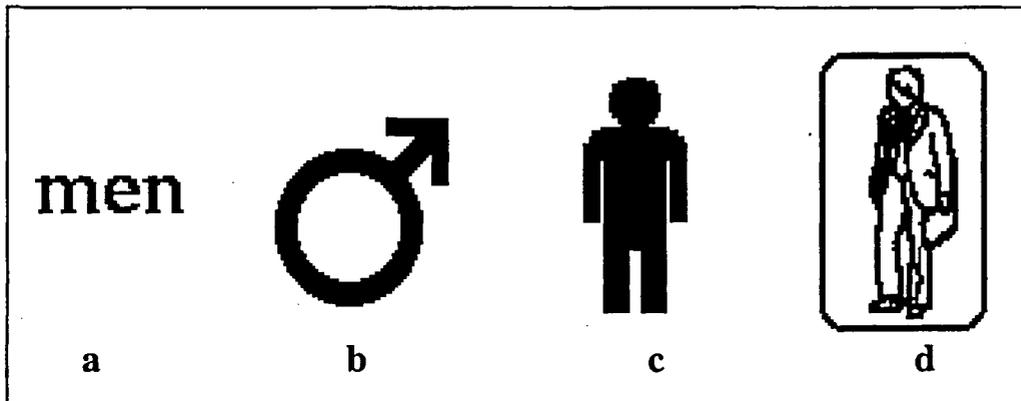


Figure 1: Different ways of referring to 'men'

Words are essentially arbitrary in the way they refer, so there is no alternative but to learn their meaning. Of course words may be morphologically complex (*spaceship*, for example), but the components themselves are ultimately atomic and have arbitrary referents. With icons, the relationship to the referent is not arbitrary, but neither is it as direct as in the case of pictures. It would be quite wrong to see pictures as an ideal form of icon because there are a number of things they are not good at expressing. For example, it is difficult to draw a

naturalistic picture of *love* or of general classes of things, such as *the class of all mammals*. For the purpose of pictorial language, pictures reveal too much information: the man in Figure 1(d) might stand for *businessmen*, *middle-aged men*, or any of a number of other interpretations. Quite apart from this, the amount of detail means that differentiation from other symbols is not so easy, and therefore recognition can be considerably slower than with the more stylized forms.

While an icon suggests its referent, its form is often insufficient to describe it precisely. What advantage, then, does it give us over words? There are two reasons for which an icon can be

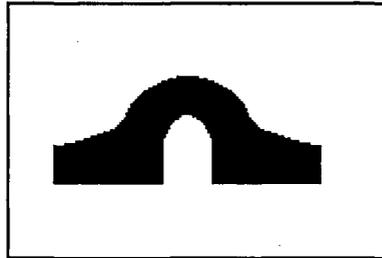


Figure 2: A hump-backed bridge

advantageous: it can be easier to learn and easier to remember. It can be easier to learn because its appearance at least suggests a set of possible referents and because it is often part of a consistent system of representation which itself will provide a context. For example, the symbol in Figure 2 is given its reference by being placed on a white background in a red triangle on a black and white striped poll beside a road. This tells us immediately that it is a sign intended for drivers and concerns some potential hazard. It can be easier to remember because it is a well-known aid to memory to associate the thing to be remembered with a simple object to which it has a defined relationship. For example, to remember the French word for *pig* (*cochon*)

I may think that it sounds like the English word *cushion*. I therefore associate the concept of *the French for pig* with an image of a cushion, and this serves as an aid to memory.² This illustrates well the fact that an icon does not necessarily have to be directly representational but can indicate its reference by means of a convention, as in Figure 1(b). Such conventions are of course also applicable to user-interfaces, including those in CAL programs.

Iconic languages

It has been suggested that most modern written languages are derived from pictorial languages, but iconic languages seem to be a more recent invention. Some of the more interesting examples are Semantography (Bliss, 1965), Isotype (International System Of Typographic Picture Education) (Neurath, 1978) and Worldsign (Jones and Cregan, 1986), a language created for mentally handicapped children which allows dynamic representation.

In Semantography (or Bliss symbols), one symbol corresponds to one word in natural language. It was proposed as an auxiliary writing tool for communication between different nations, and as a device to specify relative and vague meanings. The aim of the language is to communicate through simple pictorial symbols, with those representing physical things using outline and those representing non-physical things using geometric symbols. The first 25 symbols are already internationally accepted; they include the digits 0 to 9 and symbols such as a question mark, a full stop and a plus sign.

Bliss based his grammar on the assumption that all languages are used to describe the phenomena of our physical world, that the main manifestations of our world can be classified into *matter*, *energy* and the *mental*, and that everything happens in space and time. There is a

specific logic behind symbols constructed to represent words. A word like *telephone* takes its symbolization from the symbols for *electric* and *language* (a telephone is an electrical apparatus in which you can talk) and the *language* symbol is derived from the *ear* and *mouth* symbols (both used in conversation).

The Isotype picture language does not have a simple correspondence between signs and words. An example given by Neurath is that there is no sign for the word *foot* that is common to expressions such as *the foot of a man*, *the foot of mountain*, and *the foot of a table*. These expressions are composed of simple signs of a very different sort. Furthermore, the final translation of the 'language picture' is a structured group of statements, and the system of connection between signs is far richer than in linear text. Several rows of connected signs are interpreted simultaneously, whereas one-dimensional text requires readers to bear in mind what they have read and to make connections between dispersed elements for themselves.

Neurath's writing suggests two central rules for generating the vocabulary of an international picture language: reduction, for determining the style of individual signs, and consistency, for giving a group of signs the appearance of a coherent system. Then there needs to be a set of conventions to allow the user to know how the information is structured. Neurath's work was directed towards making statistical charts. He introduced two basic rules: the first of these related to the presentation of statistics by means of icons, and held that a icon represents a certain quantity or amount of things and that more signs represent a greater quantity or amount. The second was a general rule that perspective should not be used. His work includes a series of posters for an anti-tuberculosis campaign (non-statistical) and the publication of many books and charts.

The Bliss approach is similar to text-based language and susceptible to some of the same pitfalls, but it presents some interesting insights for an international human-human computer system. Isotype in its present form is best suited for statistics and is too restricted for a computer environment but it could be seen as a precursor of an iconic computer language.

Computer-based icons

The computer provides a significant new environment for the use of icons. Whereas previous systems of icons have been essentially textual, their messages are one-dimensional sequences of icons in which everything is explicit. Each icon has its reference, and icons are placed in sequences (occasionally there are sub- and super-scripts) with blank spaces to indicate grouping.

The typical windows environment on a modern computer provides a much richer environment. First, there are fully two dimensions to exploit so that any grammar which exploits relative positioning of icons is not restricted in the way that text-based languages have been (though from a practical point of view there are limits on how much one can present at one time). There is also the possibility of using colour (or greyscale) and animation.

Most importantly, an icon in a computer environment needs to be defined both representationally and operationally. In addition to asking what an icon represents (pictorially), one can also ask what happens when the user clicks on it, or double-clicks on it, or clicks-and-drags it to some other part of the screen. Here we are opening up a world in which the sentence 'Don't look for the meaning, look for the use' has a new potency. The

meaning of an icon is no longer simply what it resembles, but also what happens when you do something with it. This realization gives a dramatic new lease of life to what may appear to have been a marginalized form of communication.

The remainder of this paper describes some of the work we have undertaken so far in this area. It is centred around two example systems: a hotel booking system (developed at the University of Exeter) and CD-Icon (developed at Brighton Polytechnic). While the hotel booking example does not relate specifically to learning technology, the principles on which it is based are relevant to any user-interface design, and CD-Icon has obvious similar relevance.

Example 1 – Hotel booking system

Hotel booking is a typical activity that requires communication across language barriers. It offers us the opportunity to apply iconic language in a simple dialogue between a potential guest and a hotel or city-wide hotel booking facility. In a final system we can envisage a touch-sensitive screen with plenty of interaction, but at this stage of development we are concerned with the initial formulation of a request by the customer that will be transmitted to a hotel for the compilation of a reply. A demonstration system has been built using HyperCard.

The compilation of the booking message is accomplished in stages, and at each stage the current context is cued by a picture resident in the background. In sequence these are: a 'typical' hotel front (Figure 3), a 'typical' hotel reception area (Figure 4), and a 'typical' hotel bedroom (Figure 5). Each new screen holds the background picture before the other information is faded in over it.

The initial screen shows a hotel overlaid by an appropriate caption, and clicking anywhere on the image starts the booking sequence. The first screen invites the user to indicate the intended destination (the name of a town or a hotel) and grade of hotel, by selecting from cyclable 'star' ratings (Figure 3). Movement to the next screen is initiated by clicking on the 'tick' icon, a convention used throughout the package.

The second screen (Figure 4) shows a hotel reception area and invites selection of the dates and times of arrival and departure. The number of nights that are implied by these dates and times is indicated by black bars which appear (and disappear) as each night is added (or removed). Once again, the 'tick' icon moves the user to the next screen.

The third screen (Figure 5) shows a room overlaid with icons permitting the selection of rooms and their required facilities. A room is shown as a white rectangle. One room is shown initially and the number of rooms can be altered by clicking on the '+' and '-' icons. Four icons at the top right of the screen each unlock further related icons to enable the selection, for each room, of: (a) the number and type of occupants, (b) the number and type of beds, (c) the type of bathroom facilities, and (d) the range of other facilities required. The various features are selected by clicking on an icon which causes a clone to be produced beside it, which is then dragged into the relevant room. The 'tick' icon moves the user to screen 4 (Figure 6) which displays the complete booking requirement. If this is satisfactory a further 'tick' sends the message to the hotel.

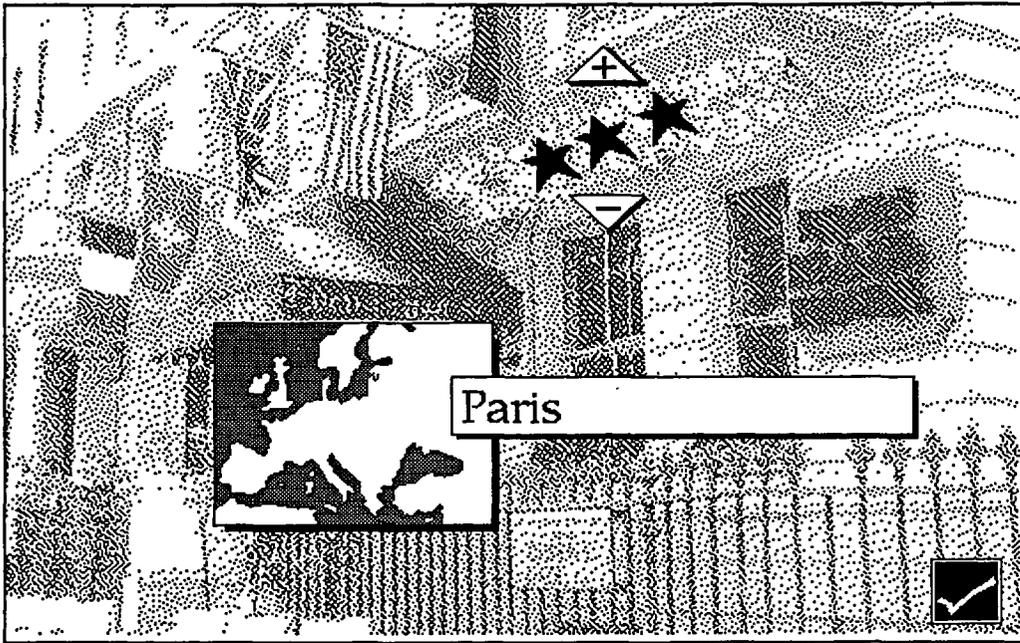


Figure 3: Screen inviting input of destination

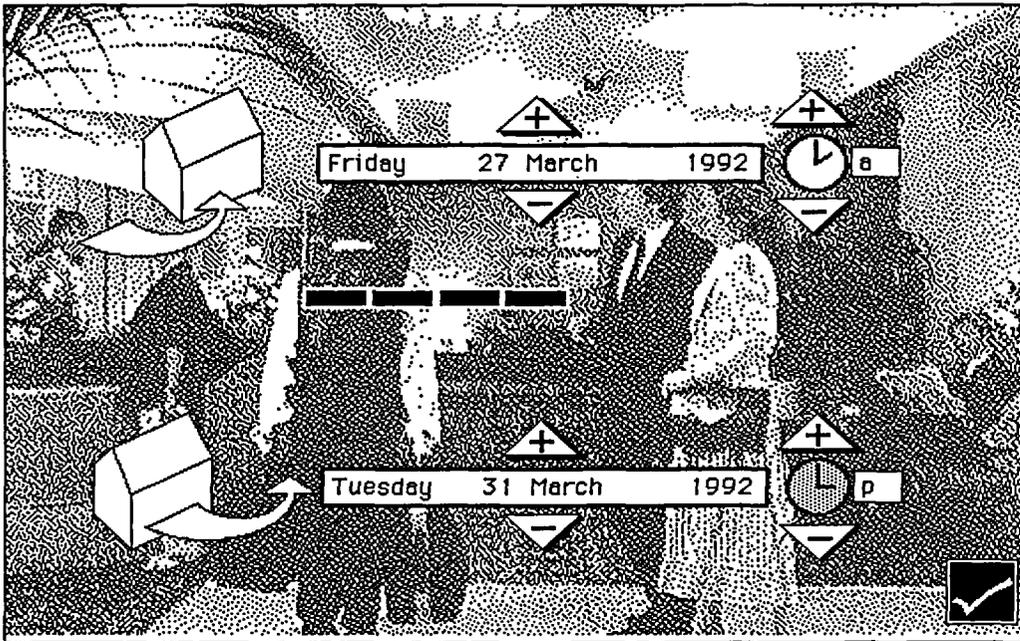


Figure 4: Screen inviting selection of dates and times of arrival and departure

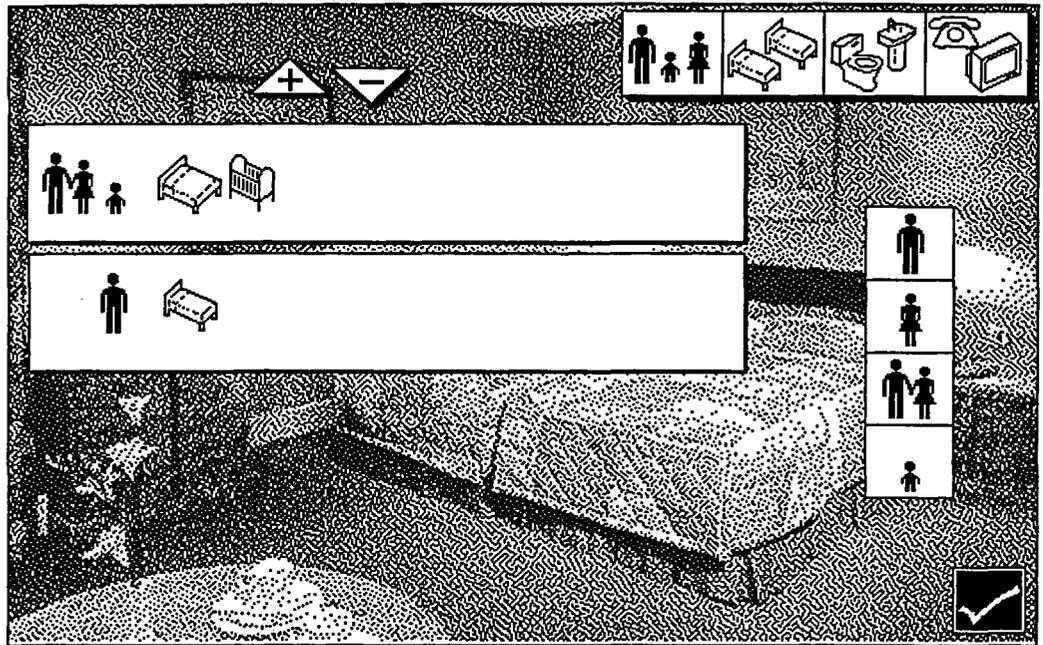


Figure 5: Screen allowing selection of room types and facilities

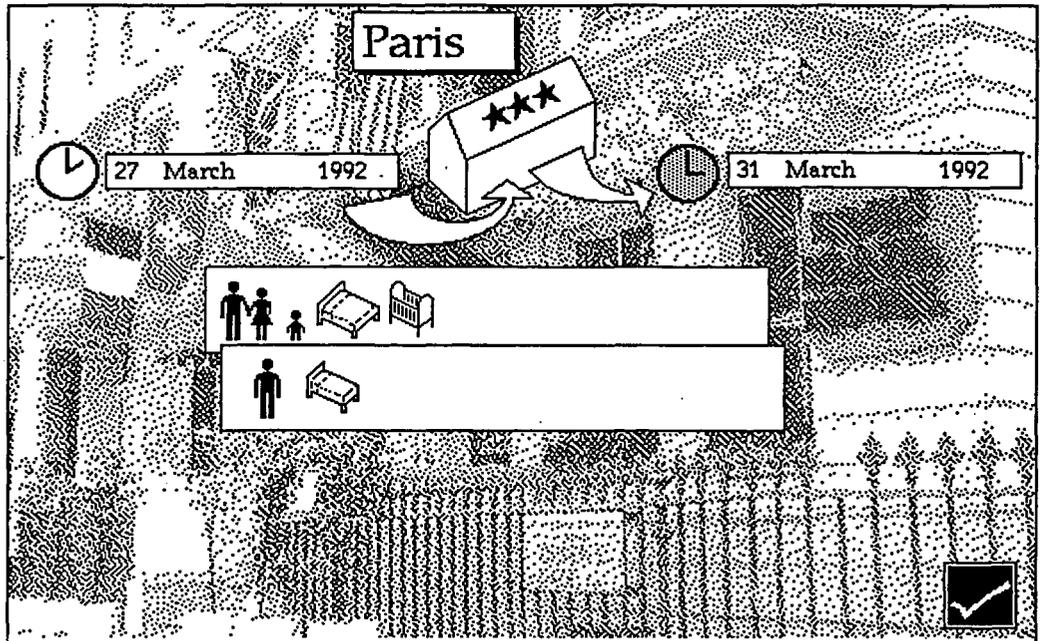


Figure 6: Screen displaying complete booking requirements

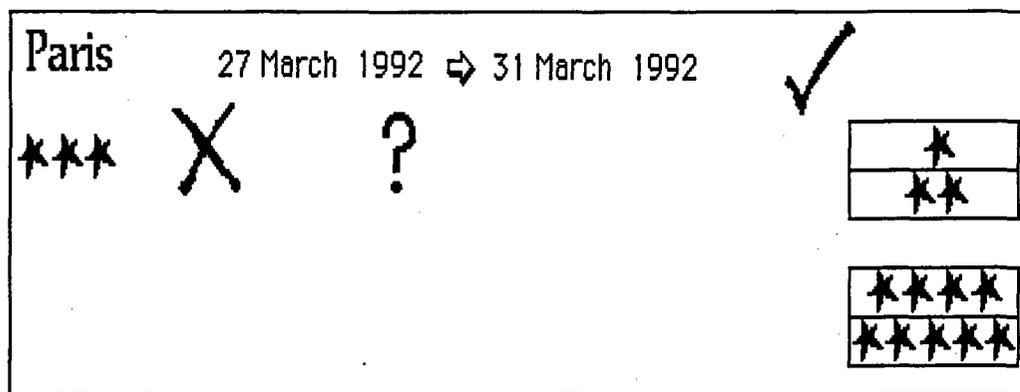


Figure 7: Screen showing availability of 'star' ratings

The message is revealed to the hotel in stages. Confirmation of the acceptability of each part of the message (tick) moves on to the next part of the message, whilst unavailability (cross) brings up a range of possible alternatives. Figure 7 shows that the required dates of stay are acceptable but the requested hotel grade is unavailable. The hotel is therefore presented with four options to propose in reply.

Once the entire message has been processed by the hotel in this way, the final message is sent back to the customer who will be able to accept or reject the alternatives offered, continue the dialogue, and confirm a booking.

The application, as it stands, does not pretend to be either comprehensive or the most practical solution in real terms, but is an initial attempt to create a simple, interactive, iconic dialogue using hotel booking as a convenient theme. It does, however, offer much that could be used in a real system, and serves its purpose in starting to explore the possibility of communicating with icons.

Example 2 – the CD-Icon language

The CD-Icon system is an attempt to build an iconic communication system based on the principles that underlie natural-language processing systems. The standard way of specifying semantics in such systems is to assume some other system (a Meaning Representation Language, or MRL) for which the semantics are already known (Woods, 1978). The task then becomes that of expressing rules by means of which statements in natural language are transformed into statements into the MRL, and vice versa. Schank's Conceptual Dependency representation (Schank, 1973) is an MRL in this sense, and its own semantics are either taken to be intuitive, or are established by their success in various practical projects, for example MARGIE (Schank, 1975) and PAM (Wilensky, 1981).

CD-Icon is a means of testing the validity of Conceptual Dependency directly by making it the basis of a communication system that uses only icons and no words. A message is composed by selecting options from a series of interconnected screens (in the spirit of

systemic grammar). The message is then transmitted, also as a set of interconnected screens, but not showing options that have not been selected.

We will illustrate the system by composing the message equivalent to 'The big man went home'. Figure 8 shows the message expressed in Schank's CD formalism.

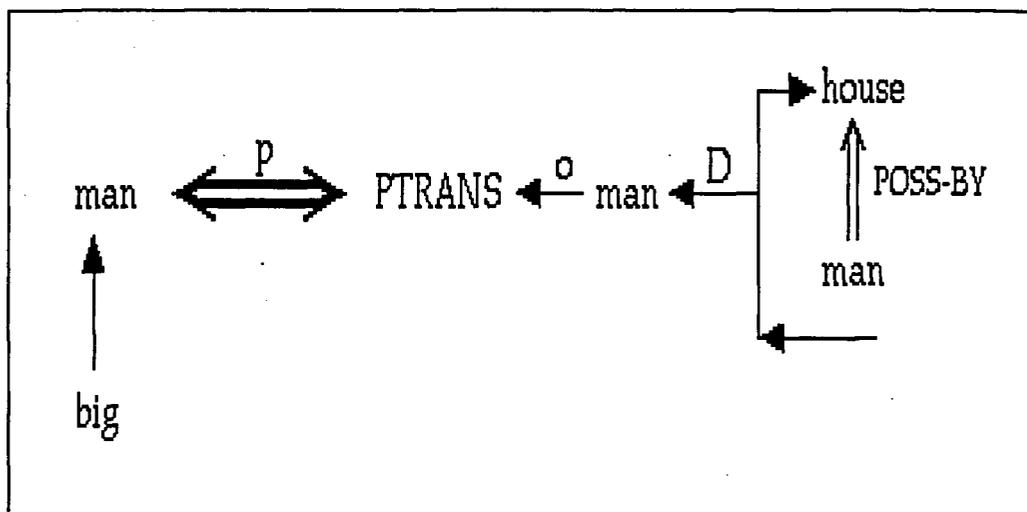


Figure 8 Schank's CD representation of 'The big man went home'

In CD-Icon a message is composed in four stages. The first stage is concerned with what Schank calls 'conceptual relations' and 'conceptual tenses'. A screen is presented (Figure 9) which enables the user to select an assertion, a question or an imperative, between simple and compound messages, and to decide upon negation. If the message is compound, the nature of the relationship between the two component conceptualizations is chosen (logical AND, logical OR, implication, temporal or spatial).

Clicking on an icon for a conceptualization transfers control to stage 2 which is typically concerned with an event. Events, according to Schank, are based around a primitive 'act' so the first selection is between icons representing various primitive 'acts' (Figure 10). To assist the user, a Help facility presents a short animated explanation, in the manner of Mealing and Yazdani (1990). Having selected the appropriate 'act', the corresponding screen is presented. In this example it is the PTRANS screen which is shown in Figure 11. It contains a basic background for PTRANS with grey icons representing the object, origin, destination and instrument cases, as well as tense. (There is a divergence from Schank here in that we do not use the agent case.) Grey icons denote options, whereas black-and-white icons represent selections. A grey house, for example, represents the class of places, the grey question mark represents the class of objects, the grey clock represents the class of times, and the grey spade represents the class of instrument cases.

Clicking on any one of these icons (except the clock or spade) will result in transfer to stage 3 which is concerned with the production of what Schank refers to as a 'picture'. The 'picture'

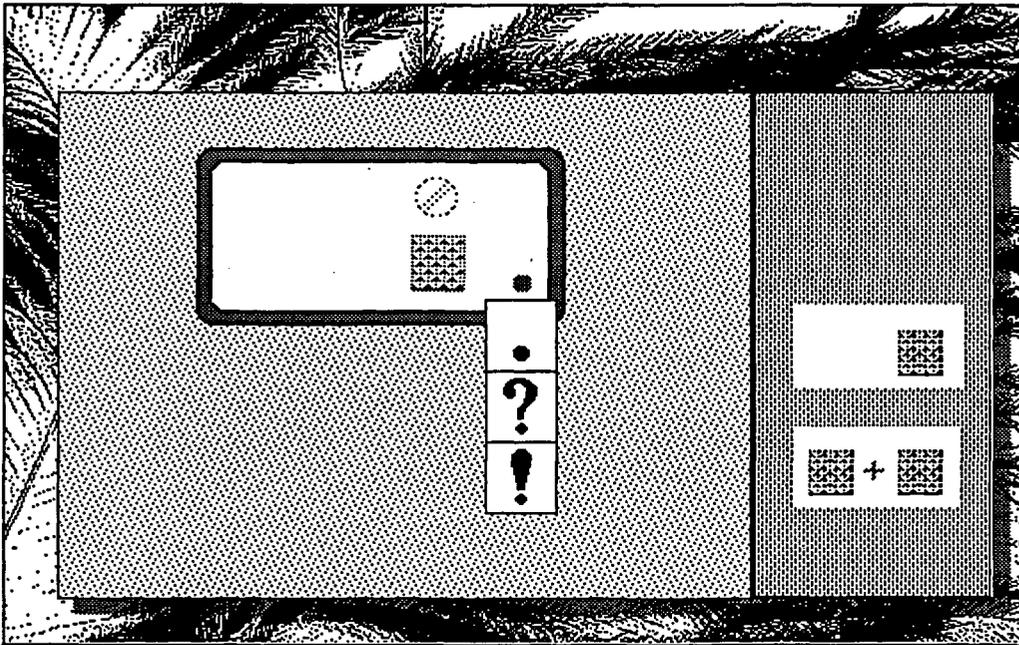


Figure 9: Screen for selecting a simple assertive sentence

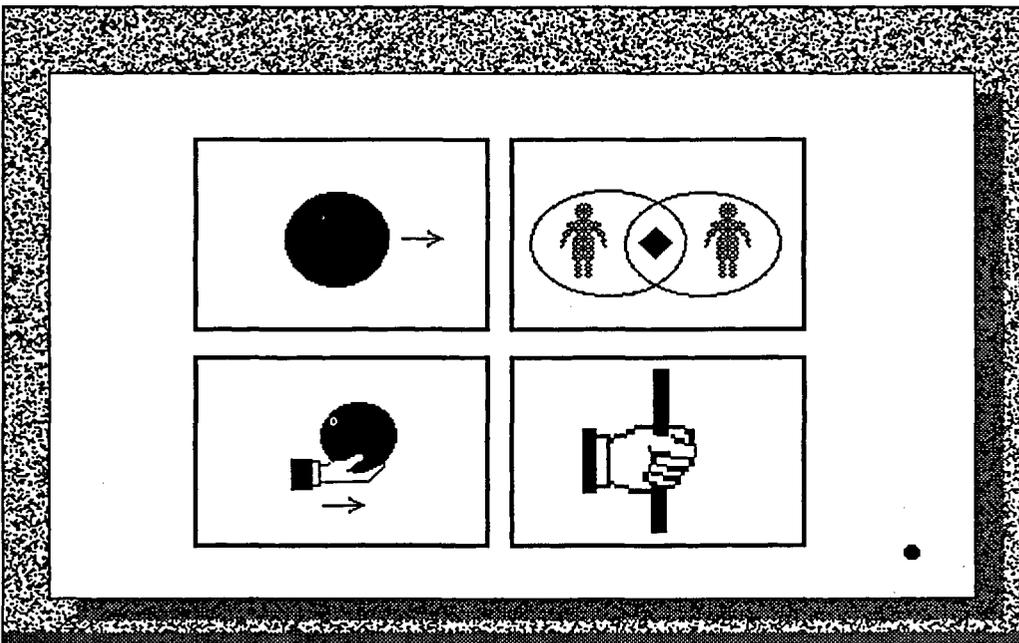


Figure 10: Selecting a primitive 'act'

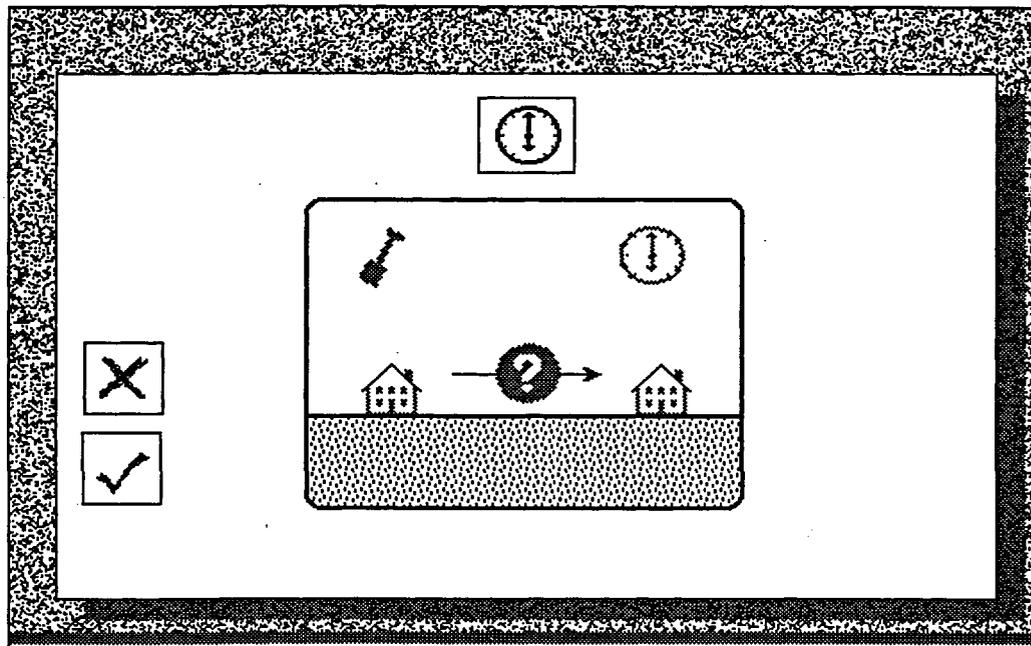


Figure 11: The screen for PTRANS

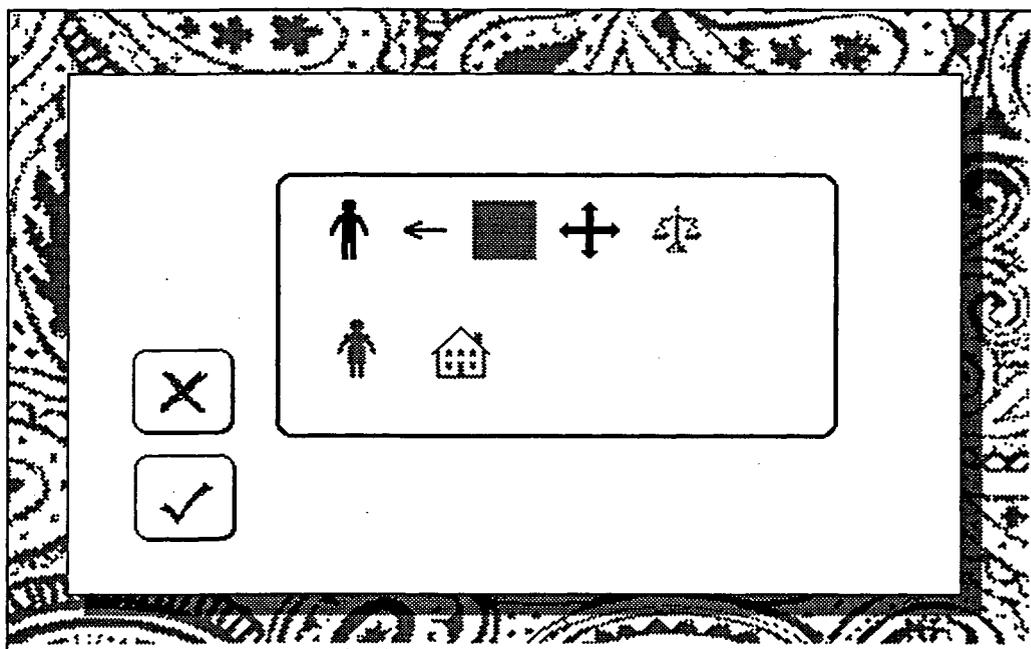


Figure 12: Picture screen for 'big man'

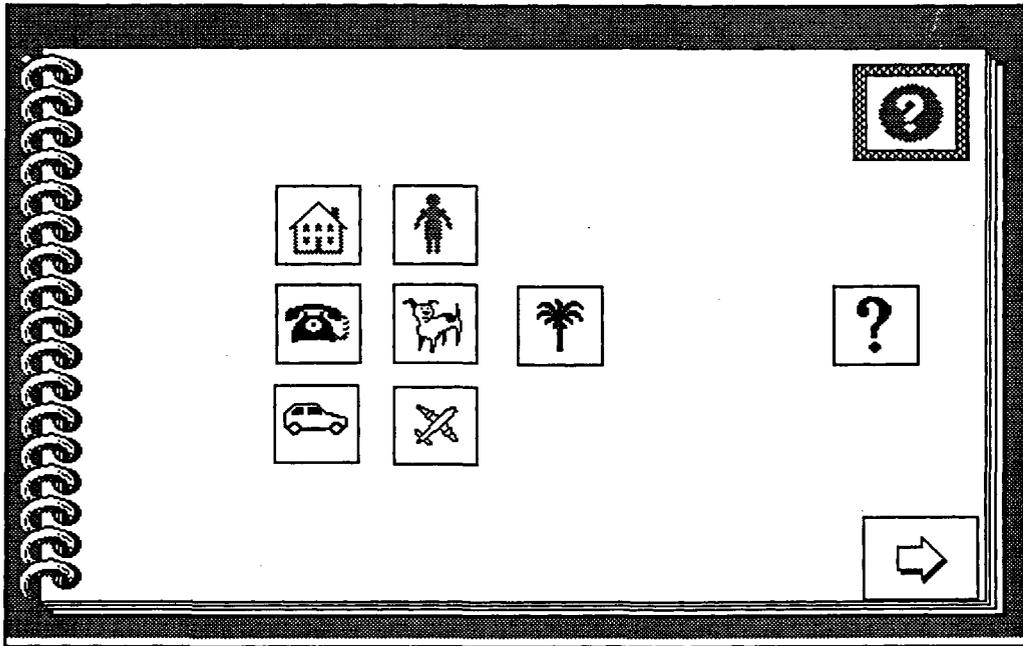


Figure 13: A screen from the lexicon

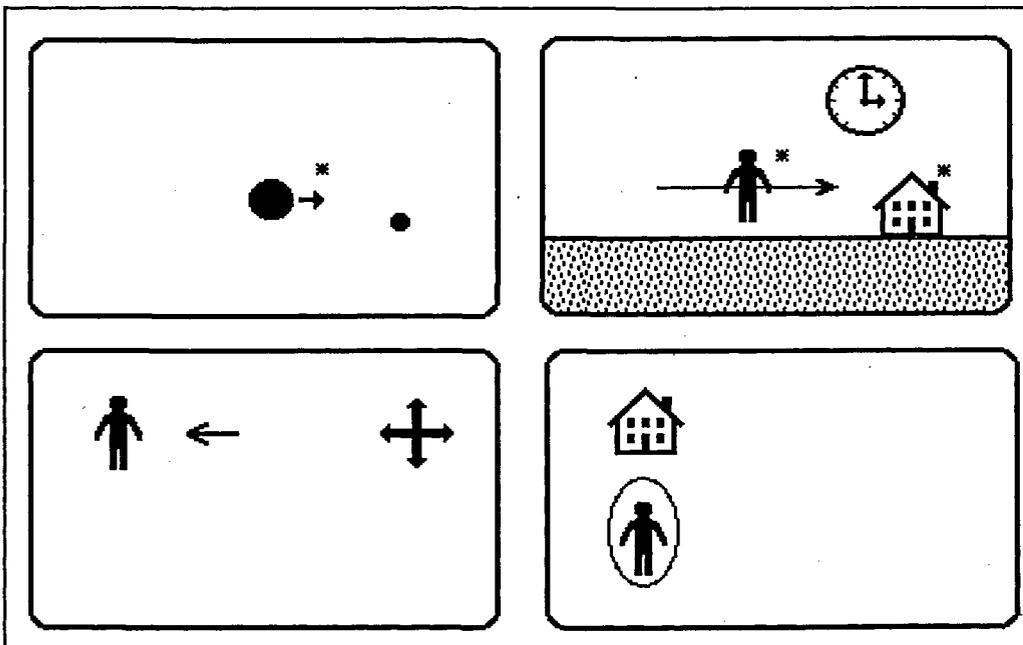


Figure 14: The four screens representing the complete message

screen initially contains only the head icon (the icon that was clicked on at stage 2). This will eventually be defined along with any modifiers to produce the screen in Figure 12.

Clicking on the 'head' icon will result in transfer to stage 4 which is the lexicon, or rather that part of the lexicon which contains objects of the type specified by the grey icon (see Figure 13). The user selects an appropriate icon, in this case the one for 'man', and control is returned to the 'picture' screen (stage 3). There are two changes however: the selected icon replaces the grey icon, and those classes of object which normally modify the 'head' icon are shown in grey.

The user can click on any of these modifying icons and be taken to the appropriate part of the lexicon to select the precise colour, size, location, etc. In our example, there is one modifier, 'big', and the outcome of stage 3 is shown in Figure 12. When the user is satisfied, the tick box is clicked and the grey icons are deleted. Control is passed back to the PTRANS screen (stage 2) with an icon composed of the head icon from the picture plus an asterisk in the top right corner. If this new icon is selected in Help mode, the full iconic representation will be displayed. The process is repeated for all icons which the user decides to specify.

Temporal reference is established with respect to an imported clock icon representing 'now' (set at 6 o'clock). A clock internal to the 'act' can be set at 'past', 'present' or 'future' (3, 6, or 9 o'clock). In our example, the past tense is used.

The instrument case is handled differently. In Schank's system the instrument case never points to an object, but always to a conceptualization. In CD-Icon, control will be passed recursively to stage 2 to devise a new conceptualization, which will be represented by a new icon (a black spade with an asterisk) by means of which the instrument case can be made explicit.

When the user is satisfied, clicking on the tick box will return control to the message-level screen (stage 1) with an icon for PTRANS plus an asterisk in the top right corner. The final message will be represented by the four interconnected screens shown in Figure 14.

At present the system is being used to explore the validity of MRLs and the possibility of unrestricted communication by icons. It is hoped to soon have a system that will allow users to try to compose and understand simple messages, at which point empirical testing will take place.

Future directions

These two projects have already raised some important issues. We need to distinguish different communicative tasks, for example to distinguish between an iconic system that serves as a front-end to existing computer software and an iconic system for person-to-person communication. This distinction seems to mirror the distinction between systems that have a known MRL and those that do not.

The systems also raise the possibility of a complete escape from linguistic forms. At present there is a tendency to explain them with reference to linguistic examples – that is to say, one tries to compose a message that corresponds to a sentence which has already been formulated in English. The intention is, however, to escape from this and view the systems as communication channels in their own right. The output should not be verbalized, except

perhaps in some indirect way in order to test the degree to which communication has taken place.

This raises the question of the form of the communication itself. In the hotel booking system it is a single screen, whereas in CD-Icon it is a set of connected screens. The question is being considered, particularly when dealing with larger texts, of whether an animated screen may not be more appropriate. In the case of our second example, the 'man' icon could appear moving to a house. This would be no ordinary animation, for the actors would be iconic and the animation may be stopped at any point and icons selected to reveal more information about themselves.

Notes

1. Cross-language communication means communication irrespective of the language spoken by the participants. It is distinguished from cross-cultural communication which raises a number of specific problems, and while we can see how cross-cultural issues might be addressed, to date little research has been carried out in the area.
2. This system of language learning has been commercially exploited (Gruneberg 1987-1992, Gruneberg and Jacobs, 1991).

References

- Adams, D. (1979), *The Hitch-Hiker's Guide to the Galaxy*, London, Pan Books.
- Bliss, C.K. (1965), *Semantography*, Australia, Semantography Publications.
- Gruneberg, M. (1987-1992); Gruneberg, M. and Jacobs, G. (1992), *Linkword Language System* (various languages), London, Corgi Books.
- Gruneberg, M. and Jacobs, G. (1991), 'In defence of Linkword', *Language Learning Journal*, 3, 25-29.
- Kobayashi, K. (1986), *Computers and Communications*, Cambridge (Mass), MIT Press.
- Mealing, S. (1992), 'Talking pictures', *Intelligent Tutoring Media*, 2, 2, 63-69.
- Mealing, S. and Yazdani, M. (1992), 'A computer-based iconic language', *Intelligent Tutoring Media*, 1, 3, 133-36.
- Neurath, O. (1978), *International Picture Language*, University of Reading, Department of Typography and Graphic Communication.
- Schank, R.C. (1973), 'Identification of conceptualisations underlying natural language' in Schank, R.C. and Colby, M.C. (eds), *Computer Models of Thought and Language*, San Francisco, Freeman, 187-247.
- Schank, R.C. (1975), *Conceptual Information Processing*, New York, North-Holland.
- Wilensky, R. (1981), 'PAM' in Schank, R. and Reisbeck, C. (eds), *Inside Computer Understanding*, New Jersey, Lawrence Erlbaum, 136-179.
- Woods, W. (1978), 'Semantics and quantification in natural language question answering' in Yovits, M. (ed), *Advances in Computers* (17, 2-64), New York, Academic Press.