DIVISION OF COMPUTER SCIENCE

A Knowledge Extraction Approach for Capturing Dispersed Knowledge and its Application to the Modelling of Working Environments for People with Disabilities

Jill Hewitt
John Sapsford-Francis

Technical Report No.194

March 1994

A Knowledge Extraction approach for capturing dispersed knowledge and its application to the modelling of working environments for people with disabilities

Jill Hewitt, John Sapsford-Francis, School of Information Sciences, University of Hertfordshire, College Lane, Hatfield, Herts. AL10 9AB ENGLAND.
Tel: +44 (0)707 284327 Fax: +44 (0) 707 284303

This paper further develops the Generic Reference Model Approach to modelling working environments [1,2], by describing a method of knowledge extraction which is suitable for use by non-expert practitioners in a variety of countries. The method, TOX, was developed to solve the specific problems encountered in the MODEMA[3] project which is building a knowledge based browsing system to support queries relating to the integration of people with disabilities into paid employment.

1.0 Introduction

This paper refers to ongoing work on the MODEMA project which is more fully described in [3]. A two-part knowledge extraction methodology was designed which would capture the depth of an expert's knowledge as well as a broader view of work environments and the current role of disabled people in them. The Task Oriented Cross-Referencing (TOX) approach which is described in this paper is being used to build the models of working environments.

2.0 The Generic Reference Model - Introduction

Work has already been carried out with British Telecom [2] to develop a multiperspective approach to modelling an office environment. The original aim of this work was to build a model of a working environment which could be used to test possible scenarios of use of new communications systems at a very early stage in their development. The relevance of this work to the MODEMA system is that the model could be used:

i) to assess the suitability of a person with a particular disability for a type of employment
ii) to hypothesise about the need for a particular type of compensatory equipment - either to choose a relevant piece of existing equipment or to establish gaps in the market for new equipment
iii) to provide a 'typical' job profile as an example for someone seeking work

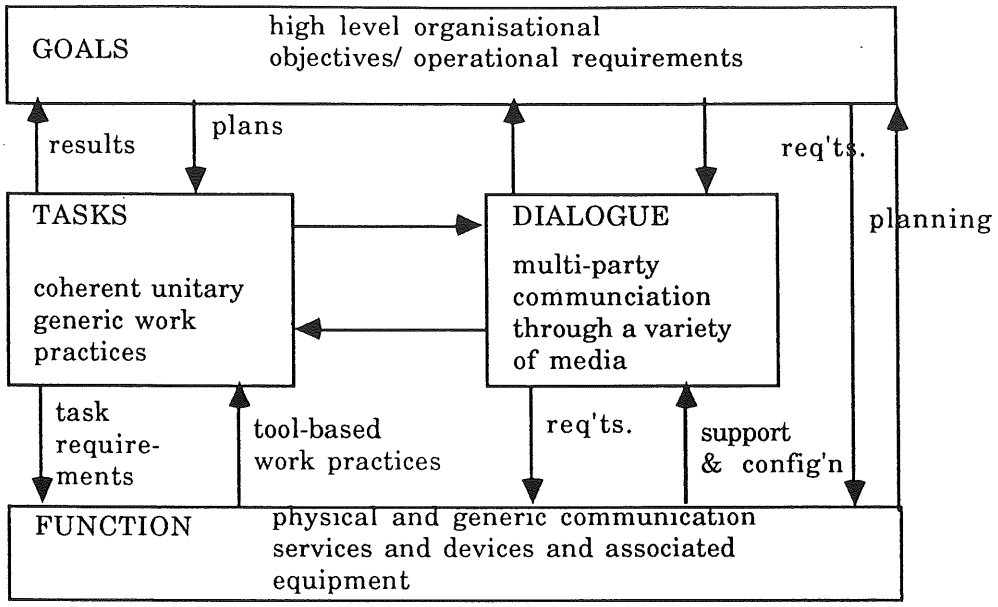
The original method of generating a model is described in the next section, but for the purposes of MODEMA, a simplified approach was developed to establish a very broad view of employment environments in different countries. The simplified method used is described in section 3.0.

2.1 Building a Generic Reference Model

From the diagram in figure 1 it can be seen that four 'views' of a working environment are considered:

i) the goal view - what does the user or organisation have to do and where do they want to get to?
ii) the task view - what tasks do people in the organisation undertake in order to achieve their goals
iii) the dialogue view -who do people in the organisation communicate with in order to carry out their tasks and what information do they need to communicate?
iv) the functional view - what existing or proposed technology is available and how does it support user activity?

Figure 1 The Office Reference Model



Together, these views form a very rich picture of the environment against which hypotheses can be measured. Typically the views are built up through semi-structured individual and group interviews with the opportunity for the users to revise the view of the environment. In addition observational techniques are used to gain a view of communication channels and typical dialogues in an office environment.

A notation - GTR [1] was developed to capture the details of the task view and the relationship between task performers and between tasks and equipment. The method draws from those developed by Johnson [4] and Diaper[5] but is designed to deal with sparse data and to incrementally build a generic task model whilst maintaining close links back to the source data. An example of a detailed task view is given in figure 2.

Figure 2 Example of detailed task view

Arrange meeting with customer: From home,car or office

GET(CUSTOMER_DETAILS, Administrator,Self)
MAKE(TELEPHONE_CALL,self,CUSTOMER)
ENTER(APPOINTMENT,self,DIARY)

2.2 Towards a simplified approach

There were problems in translating the Generic Reference Model approach as described above into a method suitable for use in the MODEMA project. We needed to gain knowledge about the organisation of working environments in order to build generic models of two environments for the prototype system - the office and the shopfloor. The method needed to capture 'normal' working practices as well as the adaptations which had been made for people with disabilities. In addition, the requirements were :

- it must be portable across international boundaries
- it must allow knowledge extraction by practitioners with varying degrees of expertise

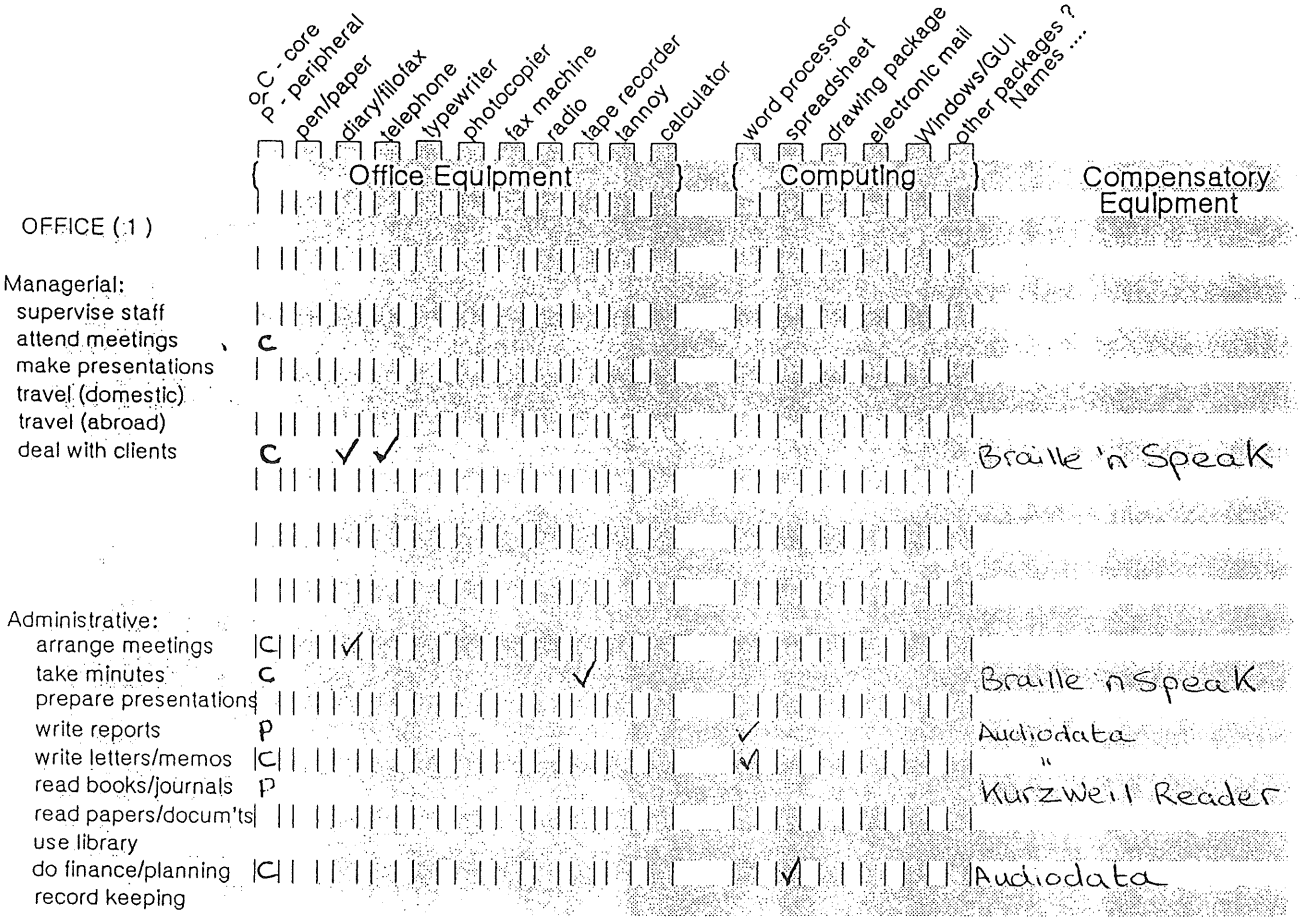
- it must be possible to use it in a wide variety of settings
- it must be capable of testing existing hypotheses about the knowledge base
- it must capture knowledge in a form that is readily convertible into the form necessary for the computerised system.

After some experimentation with interview and questionnaire formats, we decided to concentrate on the task view and developed a method to capture the relationship between tasks , equipment and compensatory equipment. This would contain an implicit reference to the function view, but would be less strong on the goal and dialogue views, thus giving a good picture of what people did and what equipment they used, but less information on the overall company objectives and the communications infrastructure in the company. We considered this to be an acceptable trade-off for this project since we had already agreed that social and psychological aspects would not be adressed in the initial project phase.

3.0 Task Oriented Cross Referencing (TOX)

This method utilises an iterative approach to building a generic model and is designed to be operated over several iterations, using a checklist that has been refined where necessary to better reflect the generic model. In addition to the checklist, a short questionnaire is provided to elicit such details as job title, disability, length of employment, type of organisation etc..

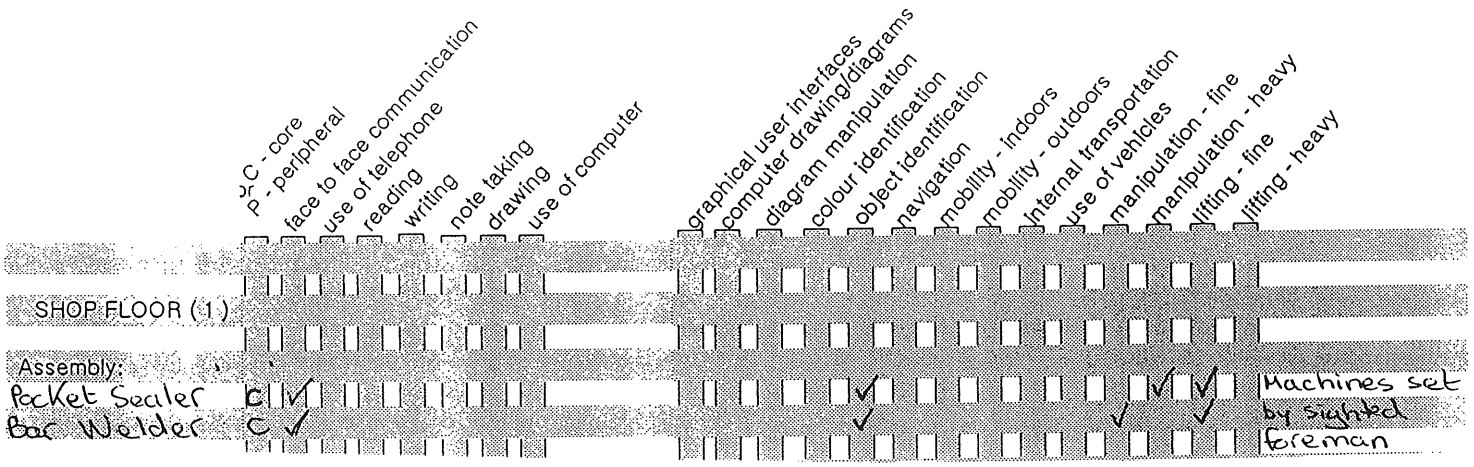
Figure 3. an extract from an office check list



A preliminary task analysis is carried out for a specified environment to develop a set of hypothesised generic objects and actions. These are represented on a "cross referenced checklist" (see figure 3) which is used in subsequent knowledge capture interviews. The left hand side of the checklist contains lists of typical actions, grouped according to the 'role' to which they belong, so for example 'supervise staff' comes under the managerial role, whereas 'take minutes' is under 'administrative'. The role headings are there for guidance, but in practice most office jobs involve tasks from a number of roles, although they tend to cluster around one or two. Along the top of the form are a list of typical tools used to carry out the tasks - these give the function view of the generic reference model, but in the case of disabled employees, this may need to be augmented by compensatory equipment. The right-hand side of the form is reserved for references to this compensatory equipment .

The example in figure 3 represents office tasks carried out by a blind employee. For example, the 'Braille 'n' Speak' notetaker is used to take minutes at meetings, and an Audiodata screen reader is used in conjunction with a word processor to write reports.

Figure 4 an extract from a shopfloor checklist



Because the use of office tools such as telephone, fax and photocopier is well understood, it was not deemed necessary to include a further description of their operating requirements. However, some tasks on the shop floor involve the use of tools and machinery which may be unique to a company or at least not commonly understood. For this reason, the checklists for the shopfloor have space to record details of required physical functions such as 'heavy lifting' or 'object identification' - see figure 4. In addition, at the time of writing the authors did not have a very clear generic model for this type of environment, so the left hand column was largely left blank so that interviewers could fill in the task names as they identified them. The next iteration of knowledge extraction will contain an initial generic model based on these tasks.

It is important that the right level of task is chosen, since too detailed a breakdown could mean that alternative methods of achieving a goal are not considered. An example is the task of 'take minutes' which can be readily carried out by a blind person with a braille notetaker, but if broken down into one possible method of 'take shorthand' and 'copy typing' would appear to be very difficult for the blind person to carry out.

3.1 The TOX process

Although the checklists form the basis of a structured interview, it was not deemed to be a good idea to simply work through them with the interviewee; experience has shown that this does not give a true picture of the tasks carried out. It was envisaged that the interviewer could fill them in as s/he conducted the interview, but this has proved to be impractical in the early stages as the interviewers are not familiar with the forms and found it too disruptive to stop the interview whilst they tried to find the

right place to make an entry. An added problem is one of language - if the form is left in English then the interviewer may have trouble interpreting it, but if it is translated there is more scope for error when the results are compiled back in England (although the fixed format should go some way to alleviating this problem).

The method which seems to be the most practical so far is for the interview to be carried out without the direct use of the form and for the interviewer to fill in the (English) form afterwards from notes or tape recordings taken at the time of the interview. Where possible, it is recommended that the form is then checked back with the employee to ensure that all the data is correct.

By the time of the evaluation point, all completed checklists should have been sent back to the 'expert' knowledge elicitors. These are then synthesised and used to build and refine the generic models of the working environments. The extracted knowledge is fed into the prototype computer system which is used to validate the model [6]. Where necessary the checklists are adjusted and updated before being redistributed for the next round of knowledge extraction. In the current project a pilot extraction and evaluation phase has already been carried out and there are three evaluation points planned before its completion in June 1993. The knowledge will be extracted from five countries and the models will be based on a minimum of 100 interviews from each country.

4.0 Scenario Generation

As mentioned above, the generic task lists are currently used in the computer system to provide guidance to users in relating compensatory equipment to tasks and in identifying ways to overcome any problems which might arise when considering a person's ability to do a job. Once the model is more complete, it will be possible to generate profiles for particular roles or jobs showing typical core tasks, for example a disabled person seeking work might want to generate a typical set of tasks for a supervisor on an assembly line so that s/he could be better armed to answer questions when s/he goes for interview. These job profiles are also useful for employers and advisors in testing scenarios of employment of people with disabilities, so that compensatory equipment requirements may be considered and tasks which need further thought can be identified. A further use is for the manufacturers of specialist equipment who need to see which tasks it might be used for and where there are still gaps in the market.

5.0 Conclusions

TOX can capture diverse knowledge from many sources in a form that is easy to compare and which will facilitate the building of generic models of working environments. The described method is currently being used by both trained and untrained analysts in five european countries. Preliminary results show that the method does fulfil its aim of capturing a broad overview of the working environments although some in-depth interviews would be needed if a full generic reference model were to be built.

Further work is planned to automate the process of building a generic model and to make this process available to non-experts so they can build models of other environments.

6.0 References

1. Hewitt, J., J. Hobson & J. Sapsford-Francis , 1990, "An Application of Task Analysis to the Development of a Generic Office Reference Model", Interact '90, Cambridge, England.
2. Watkinson, N.S., S.M. Furner, & K.R. Dickerson, 1991, "A usage reference model for office systems", Br. Telecom J. Vol. 9 No. 1 January 1991
3. Halford, P., J. Hewitt, & S. Weaver, 1992 , "Ergonomic design solutions and model interrelationships for a knowledge based browsing system for diverse user groups", this conference proceedings
4. Johnson, P., D. Diaper, 1989, Task Analysis for Knowledge Descriptions: Theory and Applications in Training in Cognitive Ergonomics. Long & Whitfield (eds.)
5. Diaper,D., 1989,Task Analysis for Human Computer Interaction, Ch. 4 Ellis Horwood.
6. Sapsford Francis, J., & Halford,P., 1992, "Validating Knowledge Models across Multinational Boundaries", this conference proceedings