An Object-Oriented Case Management System for CAPI Surveys

Vesa Kuusela, Anssi Parviainen, Statistics Finland

1. Introduction

Case management is one of the most important parts of a CAPI information system. It is also one of the most difficult ones in case something more than a trivial design is required. It is Important since it deals with just what the system is all about, i.e. the collection and transferring data concerning sample points (which may be persons, households, companies etc.). The most important part, of course, is the CAI-software but that is usually a package, such as Blaise, which no longer needs any specific designing. Case management is difficult since, in order to be functional, it involves some fairly complicated manoeuvres with the cases - and since it has to be faultless.

By case management we mean the part of the CAPI information system that includes the distribution of sample points to interviewers, their delivery to interviewers' laptops, transfers of cases from one interviewer to another, collection of the sample points back to the office into a single file, and in panel surveys, the transfer of the sample points to the next wave.

In some papers case management has had a somewhat wider meaning than the one described above (see e.g. Nicholls and Kindell, 1993). However, it is easier to analyse and design the CAPI information system if the case management system is considered an inherent but separate part of the two user interfaces : the interviewer interfaces in interviewers' laptops; and the survey management interface or supervisors' interface at the office. There are also several other tasks besides case management embedded in the user interfaces both in the office system and in the laptop system. User interface provides many activities that support the case management, e.g. telecommunications and launching interviews, etc. We conceive the user interface rather as a platform where case management may be set.

Ideally the case management system should be a transparent part in a transparent interviewer interface which should facilitate the interviewer's work rather than makes it more difficult.

Most of the difficulties in the design of the case management system arises from the fact that the sample data base has to be distributed to interviewers' laptops and later the interviewer files have to be collected

into one single data base again. This is one of the major distinctions compared to the CATI system where only one database is needed.

The distribution of the sampled cases has to be exhaustive and it should be exclusive, i.e. all cases are forwarded to one of the interviewers and no case should be available for two (or more) interviewers. However, it is not absolutely necessary that the data bases in interviewers' laptops are exclusive as long as interviewers know which sample point they should interview. However, many annoyances can be avoided if the distribution of the data base is exclusive.

There is still another aspect which puts extra demands on the CAPI system: Survey organisations have quite different ways of undertaking surveys in practice. In addition, the number of different surveys for which the system must be prepared (and of course, their size and schedule) vary. Some organisations may be set up to only few (maybe only one) surveys and the sample will not change too often. In the other end are those organisations which have several surveys going on with different sampling schemes most of the time and some surveys may have very tight deadlines.

The latter of the two 'opposite' survey environments mentioned above naturally sets more stringent requirements for the design of the system. In that case the system must be very versatile in order to enable the handling many different surveys. The installation of new surveys has to be easy and the completed interviews should be available for further processing as soon as possible. The system must stand for high standards of reliability, however. In any case in a case management system there should not exist a danger of losing data. On the other hand, if the information system is designed for a lively survey organisation it can be easily adopted in less lively ones, too.

Roughly speaking there are two major approaches to design the case management system : a data base -oriented information system and an object-oriented system.

In the data base -oriented approach a file, part of the master data base, containing all sample points for a particular interviewer is sent to his/her laptop. All interviews are collected into one data base. In a trivial case the entire interview data base is sent back to the office, maybe several times. In a more sophisticated system the completed interviews are selected from the data base and they are sent only once.

In the object-oriented system each sample point is presented as a single object, which means that all necessary information to conduct an interview is encapsulated in one file. An object contains a unique identification, the name of the survey (i.e. the method), status, sample data to contact the interviewee or household, additional data of the sample point to be used in the questionnaire, messages, etc. When the object is returned it includes updated data (e.g. status) and the answers to the questionnaire attached. The object may be a message, or a questionnaire, or something else.

The object-oriented approach provides many assets compared to data base -based approach as James Gray (1995) has put forward recently

(see also Rumbaugh, et. al, 1991). For instance the assignment of cases to interviewers becomes fairly straightforward and there is no danger that the same sample point goes to two interviewers ; and case transfers from one interviewer to another may be handled reliably. If one data file is corrupted only one case is lost in the object based system. In a corresponding situation in the data base system all interviews completed so far will probably be lost.

Once the object oriented approach is adopted it also makes many other activities rather simple. For instance, by the same telecommunications session several different objects may be sent. Each object, e.g. message, new questionnaire, a batch job, etc. will call for a specific method. Moreover, in the office system, each case may be processed separately in a so-called 'flow basis' (see Gray and Anderson, 1996). In addition, new objects and new methods are easy to install.

2. The Case Management system

The information system for CAPI surveys has been constructed gradually at Statistics Finland. The first version and the principles of the interviewer interface were described in IBUC'95 (see Kuusela, 1995). The case management system in the first version was, however, rather trivial: all interviewers had a copy of the entire sample file. The system was cumbersome and strained telecommunications. The conversion of surveys with large samples, such as Labour Force Survey, to CAPI made it necessary to design a more sophisticated system.

At Statistics Finland, an interviewer may have up to ten different surveys active simultaneously with quite a different production scheme. Some of the surveys share the same sample and some have a totally different sample. It is fairly usual, for instance, that after completion of an interview of the LFS the interviewer may have to do the interviews of a barometer survey and maybe still a third (very short) interview. At the same time the interviewer may have the household budget survey active and that has a separate sample as most of the continuous surveys do. In addition, there may still be some ad hoc surveys, with separate samples. The case management system in this situation presents a complicated designing task.

At the beginning of the design both approaches were scrutinised but quite soon, however, the object oriented approach was adopted. The reason was that the new system was expected to be both sturdy and flexible. The object based approach was considered to offer more bricks to build the system. In addition, it was fairly easy to implement the system in the existing interviewer interface. The greatest modification in the user interface was the change of the programming language: previously it was programmed by QuickBasic and now by Visual Basic for Dos. The idea is that the next version of the user interface will be done by Visual Basic for Windows. The supervisory system had to be made totally anew.

The basic idea is that a sample point is a single object that carries with it all necessary information needed in an interview. The attributes of an object may include e.g. information needed to contact a person e.g. name,

address, phone number, comments from previous interviews etc.; the name of the questionnaire to be interviewed and/or answers to interviews ; and possible external background data needed in the questionnaire, e.g. register data and/or data from a previous wave.

Each object has a unique identification code, a four digit diary code. A sample point which is to be interviewed for two surveys will yield two objects and hence two different identification codes. Each object makes a separate file and the diary code is included in the name of the file. The attributes of the object are fields in the file and they are separated by special characters. Different classes are indicated by the extension of the file name. Different classes of course have different attributes.

The name of each file carries some information of the case, just like an envelope carries the information to whom the letter should be delivered. For instance, if the file contains an interview the extension of file name indicates the status of the case and its panel rotation number. This naming convention is very useful because it enables the delivery of the cases to right places without opening them. In addition, some statistics for monitoring the field work may be produced on the basis of the file names, as well.

03. Interviewing

The installation of a new survey is ‘automatic’, i.e. when the interface software finds a specific file (questionnaire object) in the input stream it launches the method which performs the required tasks to initiate the survey. This happens right after telecommunications without a specific prompt. The installation of a new survey is told by a message to the interviewer.

The processing of sample points produces a list of all cases for the survey and external file. In the object there is the external data needed and the Extern -command to perform the indexing of the external file.

The interviewer has to select the survey first, of course. When the survey has been selected the sample of the survey is displayed. The sample points are sorted by the status code of the case.

The status codes and their meanings are as follows

New

The sample point is still intact.

Unfinished

Interview started but not completed

Non-response

The case has been marked as non-responding and copied to the out queue. A non-response case may still be interviewed later.

Interview completed

Treatment of the case was completed and the case has been sent to the central unit to be forwarded to subsequent processing. It is not possible to open the case anymore except by a case specific key which may be attained from supervisors upon request. However, it is possible to browse the answers in the training interview session.

Transfer in

Previously, the case has been assigned to another interviewer who was not able to interview it. No interviewing was done.

Transfer out

Transferred to another interviewer. The sample point has been sent to the central unit to be forwarded to another interviewer.

Returned

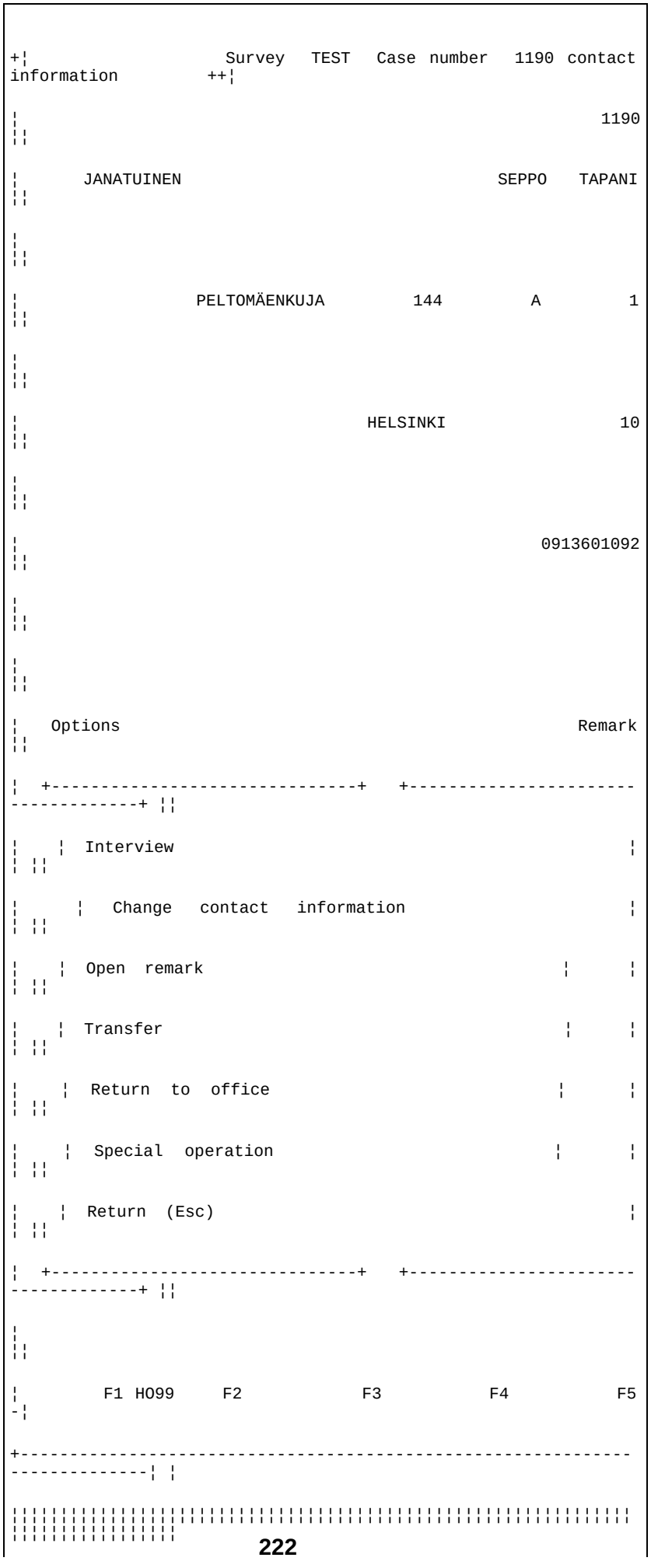
Interviewers may return a case by this option if it is unclear what to do. Option may be used only by a permission from supervisors. A returned case must have an attached message.

Special operation

Special operation is meant for exceptional situations. For instance, when the sample point for some reason does not belong to the sample.

All cases which have been sent to an interviewer have status new (or transfer in). When the interviewer closes an interview she/he has to mark the new status of the interview. It can be either unfinished, interview completed or non-response. Rest of the status codes are used in exceptional case movements. If the status was changed either to completed interview, transfer out, returned, or special operation, the case is moved to send queue and the interviewer cannot open it anymore except by a case specific key which may be attained from supervisors upon request. They may still open the non-response cases later on although the case had been sent to the office.

Opening a case for any operation starts with a display of the sample information of the case (see figure below). In the screen the Interviewer can change most of the fields (address, phone number, etc.) and attach a remark to the case. All this information will follow the case as long as the case is active in the system.


Survey TEST Case number 1190 contact information
1190
JANATUINEN SEPPO TAPANI
PELTOMÄENKUJA 144 A 1
HELSINKI 10
0913601092
Options Remark
Interview
Change contact information
Open remark
Transfer
Return to office
Special operation
Return (Esc)
F1 HO99 F2 F3 F4 F5

4. Discussion

Ideally the case management in a CAPI system might include roughly the same functionality as in a CATI system. The techniques of to-day does not make it possible, however. Another aspect is that the CAPI and CATI organisations may be so different that it is not reasonable to try to design a similar system.

The data base -oriented information system is a natural choice in CATI system because every work station is uninterruptedly connected to the file server. In a CAPI system there is no direct and continuous connection to the master file system in the office. The data base is distributed to interviewers' laptops and the connection to the office is based on telecommunications at long intervals. The object based system provides a more firm basis for design and more possibilities, at the same time.

The basic idea in the object oriented case management system is that each sample point is stored in a separate file and the interview is stored in a separate file. The object based information system is a challenging design and programming task. Maybe the same functionality could be achieved by the data base -oriented information system but it would probably be a more demanding task. And should the system be as sturdy and flexible it would a far more demanding task.

When each interview is a single object (and a file) the complete and exhaustive distribution is quite straightforward. That is manageable even in a data base based system but dealing with transfers and panels (and transfers in and between panels) are far more easy to accomplish in an object based system and it will be far more reliable.

An object based case management system is easier to install if the sample points of the survey are known in advance. This can be accomplished even in a household survey if the households are identified by a (sampled) person. This makes the use of a unique case identifications fairly easy and hence it is possible to keep track of the movements of an object throughout a panel survey. Assigning a unique identification code to each object when they are brought forth during an interview session is a somewhat more difficult task but manageable for instance by allowing each interviewer to produce identification codes within a specified range.

The object based case management system has been in use for some months now and the basic design has proved to be the right choice. It enables implementation on many new options both in the user interface and in the case management. At the moment, only Blaise 2.4 is used. Next task is to implement also Blaise III in the system so that both Blaise versions may be used simultaneously. Surprisingly, it seems that Blaise 2.x gives better support to object-based system design than Blaise III.

The requirements a case management system has to fulfil are rather common to all CAPI information systems. Therefore, especially the design of a case management system for the interviewer laptops may be

regarded as a general purpose module attached to CAI packages, whereas, the user interfaces may have entirely different requirements in different survey organisations and it is probably more difficult to find a common core for those.

References

Gray J : An Object Based Approach for the Handling of Survey Data. In Kuusela V. (Ed.) : *Essays on Blaise 1995*. Statistics Finland, 1995.

Gray J., Anderson S. : The Data Pipeline - Processing Survey Data on a Flow Basis. In Banks R., Fairgrieve J., Gerrard L., Orchard T., Payne C., Westlake A. (Eds.) : *Survey and Statistical Computing, 1996*. Association for Survey Computing, 1996.

Kuusela V. Interviewer Interface of the CAPI-system of Statistics Finland. In Kuusela V. (ed.): *Essays on Blaise 1995*. Statistics Finland, 1995.

Nicholls W.L., Kindell K.K. : Case management and communications for Computer Assisted Interviewing. *Journal of Official Statistics*. 1993, 9, pp. 623-639.

Rumbaum J., Blaha M., Premerlani W., Eddy F., Lorensen W. : Object-Oriented Modelling and Design. Prentice Hall, 1991