IMPROVING THE USE OF VISUAL INTERACTIVE SIMULATION AS A KNOWLEDGE ELICITATION TOOL

Professor Stewart Robinson
Mr. Ernie Lee
Warwick Business School
University of Warwick
Coventry, CV4 7AL
United Kingdom
Stewart.Robinson@warwick.ac.uk
Ernie.Lee02@wbs.ac.uk

Professor John S. Edwards
Aston Business School
Aston University
Birmingham, B4 7ET
United Kingdom
J.S.Edwards@aston.ac.uk

KEYWORDS
Visual Interactive Simulation, Knowledge Elicitation

ABSTRACT

Knowledge elicitation is a well-known bottleneck in the production of knowledge-based systems (KBS). This is mainly due to the tacit property of knowledge, which makes it difficult to be explicated, and hence analysed. Past research has shown that visual interactive simulation (VIS) could effectively be used to elicit episodic knowledge that is appropriate for machine learning (an AI capability that includes inductive learning) purposes, with a view of building a KBS. Nonetheless, the VIS-based elicitation process still has much room for improvement.

Based in the Ford Dagenham Engine Assembly Plant, a research project is being undertaken to investigate the individual/joint effects of visual display level and mode of problem case generation on the elicitation process. This paper looks at the methodology employed and some issues that have been encountered to date.

INTRODUCTION

Knowledge acquisition is a crucial stage in the development of a knowledge-based system (KBS). As a process, it involves eliciting the domain knowledge that a human decision-maker uses when solving a particular problem (knowledge elicitation), organising and structuring the elicited knowledge (knowledge representation), and then codifying the knowledge into an appropriate machine-executable format (knowledge execution). In spite of the technological progress in specialised tools designed to support the building of KBS, knowledge acquisition, especially knowledge elicitation, remains a well-known bottleneck. This is because the domain knowledge is usually qualitative in nature, which includes judgment, insight and informed opinions that are built on years of experience. As such knowledge is mainly tacit, the decision-makers may find it hard to verbalise. Consequently, their inability to explicate their thinking processes renders any attempt to elicit, and therefore acquire, the domain knowledge very difficult.

To date, various methods of eliciting tacit knowledge and expressing it into reader-friendly formats have been developed. These range from manual methods of interview and protocol analysis to machine-based methods like repertory grid and machine learning. Robinson et al. (2005) looked at linking Visual Interactive Simulation (VIS) and machine learning tools to improve unplanned maintenance operations in the Ford Bridgend Engine Assembly Plant. The authors showed that VIS could be used to elicit episodic knowledge in the form of decision cases that are appropriate for subsequent machine learning (a knowledge representation technique) purposes. Nonetheless, there are still issues that need to be explored, including how to exploit the functions of a commercial VIS package more fully to improve the quality and quantity of elicited decision cases. Specifically, the individual/joint effects of visual display level and problem case generation mode on the elicitation process need to be investigated.

This paper begins with the background to the research, including a discussion on the issues that help define the research aim and hypotheses. Then, the case study at Ford Dagenham Engine Assembly Plant which provides the research setting is briefly described. Following this, the methodology used to explore how VIS could be employed as an effective and efficient knowledge elicitation tool, with a view of building a KBS, is detailed. Finally, the paper concludes with an update of the project’s progress.

BACKGROUND

Robinson et al. (2005) carried out a collaborative study with Ford Motor Company (Ford) to investigate if VIS could be used to elicit decision cases from maintenance supervisors, such that they were good enough for learning the supervisors’ decision-making strategies. The aim of their study was to devise a VIS-based means for identifying and improving human decision-making. The result of this study was the Knowledge Based Improvement (KBI) methodology (Figure 1).
The KBI methodology consists of 5 stages:

- **Stage 1**: Understanding the decision-making process
- **Stage 2**: Data collection
- **Stage 3**: Determining the decision-makers’ decision-making strategies
- **Stage 4**: Determining the consequences of the decision-making strategies
- **Stage 5**: Seeking improvements

Stage 1 requires an investigation of the system and the nature of the decision-making processes. In particular, the aim is to identify the types of decisions that are taken (the decision variables) and the factors that are taken into account in making these decisions (the decision attributes). A VIS game model is then created which represents the system under study and the decision attributes. The decision-maker interacts with the VIS game, entering decisions at appropriate points. In doing so, a series of decision cases are collected (stage 2).

In stage 3, artificial intelligence tools (e.g. a rule-based expert system) are used to learn the decision-making strategies from the decision cases. The tool is then linked with the VIS game and used in place of the decision-makers (stage 4). This enables the performance of the decision-making strategies to be tested over a long simulation run. Finally, in stage 5, ideas for improving the decision-making strategies are proposed and tested with the VIS model.

In the conclusion to their paper, Robinson *et al.* identified a number of strengths and weaknesses with the KBI approach. Among the strengths were the speed of data collection, the ability to capture all attributes of the model at a decision-point and the control over the conditions of the model. In terms of weaknesses the authors recognised that human decision-makers may take less realistic decisions in a simulated environment, as they are quite likely to take greater risks when there are no real consequences to their decisions. In addition, the authors also recognised that human decision-makers may find it a very laborious and time-consuming experience to provide a full set of data, comprising of a very large number of useful decision cases. It is these issues that provide the basis for the present research, where the aim is to investigate whether and how VIS could be used as an effective and efficient means to elicit decision cases.

In VIS-based knowledge elicitation, two levels of effectiveness can be conceived. First, the decisions made in the elicited decision cases are expected to bear close resemblance to those that a decision-maker will make in equivalent, but real situations (primary effectiveness). Second, the range of attributes and decisions collected in the elicited decision cases is expected to be wide enough to train meaningful decision-making models (secondary effectiveness). On the other hand, efficiency is the range of decision cases elicited in a span of time.

In the research, two specific factors are being investigated in terms of their influence on effectiveness and efficiency. The first is the level of visual display. The hypothesis is that a better visual display, with increased visual fidelity, leads to a greater level of primary effectiveness in knowledge elicitation (**Hypothesis 1**). In other words, the closer the visual display is to the real system, the closer the decisions will be to those taken in the real system. To test this hypothesis, the decision-makers will be presented with various 2-dimensional (2D), 3-dimensional (3D) and non-immersive virtual reality (VR) representations of the VIS model and asked to interact with the model. Here, a 3D representation consists of only 3-dimensional icons displayed against a plain background, with no perspective projection (where more distant objects are drawn smaller relative to those that are closer to the eye) or any other efforts at creating photo-realism. On the other hand, a non-immersive VR representation strives towards photo-realism, and the decision-maker is able to manipulate the view of the virtual environment by using a mouse.

The second factor is the mode of generating problem cases that are presented to the decision-makers. There are two modes: regular and non-regular. Under the regular mode, the model settings will be set such that it would simulate real-life operations as closely as possible. On the other hand, under the non-regular mode, the model settings will be set such that the probability of uncommon scenarios will be greater than in real life. The hypothesis is that the use of non-regular model settings to run the VIS model leads to a greater level of secondary effectiveness in knowledge elicitation (**Hypothesis 2**). In other words, non-regular model settings are expected to develop a larger variety of problem cases that would facilitate a wider range of decision cases to be elicited, and therefore a more meaningful decision-making model to be trained. To test this hypothesis, the decision-makers will be asked to...
interact with the VIS model running under the regular, and then the non-regular settings.

Finally, it is also hypothesised that the use of non-regular model settings will lead the decision-makers to produce a wider range of decision cases in a period of time (Hypothesis 3). In other words, non-regular settings will be more efficient for knowledge elicitation than regular settings.

**CASE STUDY: FORD Dagenham Engine Assembly**

Set in an engine assembly plant in Dagenham (east London), the research is based on the hot-test operations of the ‘Puma’ diesel engine assembly line. A schematic of the hot-test area is shown in Figure 2. Engines arrive from the main assembly line (Assy line ‘B’) where they are loaded onto a hot-test ‘platen’ (a metal pallet). They are then allocated to one of the 20 test cells, where they are rigged to a testing machine and run for a few minutes. Following the test, failed engines are sent to a repair loop for rectification, while engines that passed the test are sent to a finishing area (ATD). The key decision-maker here is the ‘switch operator’, who is responsible for the allocation of engines to hot-test cells. His objective is to maintain a smooth workflow with no bottlenecks, and ensure that the workload is equitably distributed among all operational hot-test cells.

**Figure 2: A Schematic of the Hot-test Area**

**RESEARCH METHODOLOGY**

The research methodology (to be) employed can be broken down into seven main stages: understanding the decision-making process; building a ‘game’ model; piloting testing the ‘game’ model; eliciting decision cases; learning the decision-makers’ decision-making strategies; validating the strategies; and analysing the elicitation, learning and validation data. These stages are described further below.

### Stage 1: Understanding The Decision-making Process (Completed)

The purpose of this stage is to lay the groundwork for the research. To begin, there is a need to gain a high-level understanding of the manufacturing system (i.e. hot-test area in the Dagenham engine plant) and the circumstances in which a decision-maker (i.e. group leader, or switch operator) is required to make decisions. In addition, the component parts (decision variables, decision options, attribute variables, and attribute levels) of the decision-making process are identified. In effect, this is akin to the ‘broad and shallow’ phase in Barrett and Edwards (1995), where the priority is to extend the breadth of background information as wide as is feasible. As no technique alone is sufficient to elicit all types of information (Coffey and Hoffman 2003, Rugg et al. 2000, 2002), several techniques that complement each other in gaining a broad overview in this stage are used. The techniques used here include document analysis, unstructured and semi-structured interview, and observation interview.

First and foremost, paper documents like plant layout and versatility charts (detailing the hot-test activities that each personnel is qualified to undertake) are reviewed to obtain a quick introduction to the hot-test area. The plant layout serves to help put the hot-test area in the context of the entire engine plant. It informs on the processes before an engine enters and after the engine leaves the hot-test area. In addition, it also provides a record of the physical entities that make up the engine plant, specifically the hot-test area. Versatility charts serve to provide information on the manpower status of the hot-test area. They inform on the type of responsibilities that hot-test personnel in every shift can undertake. This information is especially useful when the planning of data collection commences.

Next, informal interviews are conducted with the group leader and switch operators to collect any other relevant undocumented information, and then to clarify any queries that are formed after reviewing the documents and interview materials. In addition, such preliminary interviews also help to create opportunities for establishing good rapport with the decision-makers, whose co-operation in subsequent data collection exercises is crucial. Initially, unstructured interviews are used where the decision-makers are given the freedom to cover topics that they deem fit. This is because at this early stage, a researcher who is not familiar with the hot-test area will not have enough background information to ask specific questions (like in structured interviews), or even work to cover a list of topics (like in semi-structured interviews) in an interview session. After a few unstructured interview sessions, the researcher should have collated sufficient information from both documents and interview materials to form a clearer picture of the hot-test operations. At this point, a few queries might have surfaced, and these may be answered via another few rounds of semi-structured interviews.
Improving The VIS ‘Game’ Model’s Utility

To make the game as realistic as possible and hence remove any factors that might unnecessarily influence the fidelity of the decision-making process, one would expect the model to run at the same speed as the real system. However, problems might arise if real operations take a long time to complete, or if they are to be simulated for a long period of time (say, an eight-hour shift). As such, there is essentially a conflict between keeping the game faithful to the real system and maintaining the experiment’s expediency. Since access to the switch operators is limited, game faithfulness is traded off for higher experiment expediency by running the model at a pace that is much faster than real time.

As Ford’s VIS model spans the entire engine plant, it is expected to run relatively slowly. This, in turn, is expected to slow down the entire elicitation process. Hence, every effort should be made to improve the VIS model’s utility (run-speed). A possible solution is to simplify the VIS model as much as possible without compromising the model’s validity or credibility. An option for simplifying the model lies in splitting the original model into two parts, comprising a sub-model of pre-hot-test operations (Model A), and another sub-model of hot-test operations and post-hot-test operations (Model B). In this case, as Model A runs, its output data are collected and written to a data file. The contents of the data file are then used as input data when Model B is run. Nonetheless, there is a limitation to just using the raw contents of the data file as input data for Model B. As the input data to Model B is going to be an experimental factor, it may be more advisable to use the raw contents to construct a pseudo-empirical distribution of inter-arrival times that can be manipulated easily for subsequent experimentation. The distribution thus constructed is not strictly considered a genuine empirical distribution, as the data file contents itself are output data from another simulation (Model A, in this case) as opposed to historical data collected manually from the engine plant. For the research’s purpose, approximately 33,000 inter-arrival times have been collected from ten simulation runs of Model A. Each run simulated operations for a full week.

Stage 2: Building A VIS ‘Game’ Model (Completed)

Before plunging into building a model of the hot-test operations, the nature of decision-making in the system needs to be scrutinised first, in order to ensure that the model subsequently built is fit for the research purpose. Using Mintzberg’s (1973) classification of management roles, the decision-makers’ main role is determined as that of a resource allocator. In what is essentially a passive system, the decision-makers either make proactive decisions to pre-empt a bottleneck from emerging, or passive decisions to maintain status quo (decision of ‘no decision’). Hence, there is a need to build the model in such a way as to capture both types of decisions.

The concept of a VIS ‘game’ model fulfils this need. The model is a ‘game’ because like any simulation game commercially available in the market, it is expected to run passively according to the default settings until the user proactively changes the settings. To capture the pro-active decisions, the decision-makers will be shown a running ‘game’ model, and they will be allowed to intervene and stop the running ‘game’ momentarily to amend the allocation decisions as and when they deem fit. When an amendment is made, the set of new decision values and corresponding attribute values will be saved as a decision case, which will be required later for learning the decision-making strategies. To capture the passive decisions, the ‘game’ model could be set to perform background saves of current decision values and corresponding attribute values at random intervals of simulated time.

Using the information collected in Stage 1, three representations (2D, 3D and VR) of the ‘game’ model are built. In this case, as Ford already has a current and detailed model of its Dagenham engine plant developed in Witness (a VIS software), there is no need to build a new ‘game’ model from scratch. Instead, Ford’s existing model will be adopted and adapted for the purpose of the research. Since Ford’s model is still currently in use, it can be assumed that the model is current. In adapting the Ford model, extra care should be taken to ensure the ‘game’ model’s fidelity is on a what-you-see-is-what-you-get basis. That is, the ‘game’ model should provide neither more nor less information than the decision-maker is able to obtain/perceive in the real working environment. Also, in creating the different representations, great care should be taken to ensure that the availability of information in all three representations is consistent.

Improving The VIS ‘Game’ Model’s Logical Design

Next, the remaining model will be scrutinised for its logical design. This is because the original model is built for a purpose different from the research. In the original model,
there may be some operations that require human supervision and intervention in practice (such as by the switch operator), but which have been automated by pre-set rules defined by the model builder. Hence, there is a need to remove these pre-set rules that automated the switch operations, and re-establish the facility for the switch operator’s intervention/proactive decision-making efforts. Also, for data collection purposes, the ‘game’ model needs to be programmed in such a way as to be capable of recording the decision values and corresponding status of the model (attribute values) at each intervention.

Stage 3: Pilot Testing The VIS ‘Game’ Model (Completed)

Prior to rolling out the adapted ‘game’ model to collect data, it should be pilot tested. The approach to pilot testing may be adapted from those applied to questionnaires (see for instance, Saunders et al. 2003). The purpose of the pilot test is to refine the ‘game’ model so that the decision-makers will have no problems in using the ‘game’ model and there will be no problems in recording the data. In addition, it will enable the researcher to obtain some assessment of the ‘game’ model’s validity and credibility.

Stage 4: Eliciting Decision Cases (Featuring Hypothesis 1 – Partially Completed)

After pilot testing, the ‘game’ model will be formally employed to collect data under the different research experimental settings (i.e. factor-level combinations). As mentioned earlier, the research will be exploring if and how factors like visual display level (set at levels: 2D, 3D, and VR), and problem cases generation mode (set at levels: regular model settings, and non-regular model settings) will influence the effectiveness and efficiency of VIS-based knowledge elicitation. Hence, there are altogether six factor-level combinations to investigate. From Stage 1, it is known that there are eight decision-makers who are qualified to perform the switch operations and available to help. They will be used carefully in collecting data.

In this stage, Hypothesis 1 will be tested. Under Hypothesis 1, the view is that increased visual fidelity in VIS-based knowledge elicitation will improve the KBS’ fidelity to real-life decision-making (primary effectiveness). This hypothesis can be tested using a single blind evaluation test (Jones and Miles 1998), where a selection of decision cases collected earlier from every decision-maker under each factor-level combination will be presented back to them in a randomised order. Every decision-maker will then be asked to rate how strongly he agrees with each decision case.

Stage 5: Learning The Decision-makers’ Strategies (Featuring Hypothesis 3)

This stage sets out to learn resource allocation strategies from the decision cases collected earlier. They may be used either as direct references to support similar decisions or to induce rules (Turban et al. 2005). The former, called case-based reasoning (CBR), adapts solutions used to solve old problems for use in solving new ones. The latter, called inductive learning or rule-based reasoning (RB), allows the computer to examine decision cases and generate rules that can be inferred to solve problems. Though it is not an objective to compare the quality of learning managed by CBR and RB systems, it may be interesting to find out if decision cases provided by VIS-based elicitation are more suitable for CBR or RB systems.

To support the testing of Hypothesis 3, RB will be used to learn the resource allocation strategies. Under Hypothesis 3, the view is that using non-regular model settings will lead to a wider range of decision cases being produced in a period of time. Intuitively, the latter is expected to have a positive correlation with the number of rules generated. As such, by analysing the number of rules generated for each decision-maker under various factor-level combinations and the corresponding times taken to elicit the decision cases, some light might be shed on the hypothesis.

Stage 6: Validating The Decision-makers’ Strategies (Featuring Hypothesis 2)

The resource allocation strategies learnt in the preceding stage will be primarily examined using the principle employed in the Turing Test. The Turing Test is a test designed to determine whether a system exhibits intelligent behaviour (Turban et al. 2005). According to this test, a system can be considered smart only when an individual cannot identify the system, while conversing with both an unseen human being and an unseen system. As applied to the research, the relationship between decisions developed by the system and decisions developed by decision-makers will be measured. In essence, the system is treated like a black-box, where only the attributes and decisions are analysed, and the system processes are not investigated. As such, a smart system will arrive at the same decisions as the decision-makers, when both are provided with similar sets of attributes.

Hypothesis 2, where the view is that using non-regular model settings in VIS-based knowledge elicitation will improve the KBS’ ability to make meaningful decisions in more varied scenarios (secondary effectiveness), can be tested through the assessment for intelligent behaviour. Logically, a smart system is contingent on its knowledge base being developed from data that is both accurate and covers a wide spectrum. As accuracy of data has been dealt with in Hypothesis 1, and on the premise that the decision cases are indeed of high fidelity (i.e. Hypothesis 1), then simply proving that the system is smart would imply that the decision cases do cover a range that is sufficiently wide to learn the decision-making strategy.
Stage 7: Analysing The Elicitation (Stage 4), Learning (Stage 5) And Validation (Stage 6) Data

Eight sets of data will be collected from each factor-level combination in the elicitation, learning and validation stages. As the sample size is too small, appropriate small sample statistical tests and non-parametric tests are most likely to be used.

PROGRESS UPDATE

To date, Stage 1 to 3 (understanding the decision-making process, building the ‘game’ model, and pilot testing the ‘game’ model) are completed and the research is on its way to completing Stage 4 (eliciting decision cases). Investigations in Stage 1 have uncovered the following decision variables and attributes:

**Decision variables**
1. Pre-set allocation to straight, left or automatic
2. Switch test cell on/off
3. Manual override to eject engines from waiting stands
4. Allocate a specific engine to a specific test cell

**Attributes**
1. Engine type to be tested
2. Engine type currently being tested in a cell
3. Operational status of a test cell
4. The number of engines on each section of conveyor
5. The number of engines serviced by a cell operator in a period

In addition, the detailed VIS model provided by Ford has been adopted and adapted into an interactive ‘game’. As illustration, the 2D, 3D and VR ‘game’ models are shown in Figure 3, Figure 4 and Figure 5 respectively. In a typical elicitation session, the current attribute values would be displayed/represented in the model, and would be constantly updated while the model is running. As and when the attribute values warrant a change in the current decision values, the decision-makers would be able to control the model and do so by using the ‘game’ toolbar below the model. For instance, in Figure 3, the decision values of the circled model parts could be controlled by the toolbar icons. Whenever the current set of decision values is changed, the new set of decision values and corresponding attribute values will be saved as a decision case for subsequent machine learning.

Several trips have been made to the Dagenham engine plant, including one where the prototype ‘game’ was demonstrated to two decision-makers. To establish its content validity, the decision-makers were allowed to play with the prototype ‘game’ for a while, before their comments/suggestions were sought. Their feedback to the ‘game’ has so far been positive. Following this, even more trips have been made to the engine plant as part of the planned data collection process. At the time of writing, the elicitation sessions are almost completed.

**CONCLUSION**

The paper has described a methodology to study the ways by which VIS could be improved as a tool to elicit hot-test operations knowledge in the Ford Dagenham engine assembly plant. Two ways have been identified: level of visual display, and mode of generating problem cases. In addition, working definitions of improvement have also been established: primary and secondary effectiveness, and efficiency. At present, the first three stages of the methodology are completed and the fourth stage is near completion. Subsequent work will include eliciting a
complete set of decision cases (data) from each decision-maker, and analysing them.

Although the research is based on a specific case study, it is expected that certain generic lessons will be learned that would be useful to modellers in other contexts. For instance, the likely effects of visual display on decision-making quality (fidelity), and that of the range of scenarios generated on data quality (extensiveness).

ACKNOWLEDGEMENTS
This work is being funded by the UK Engineering and Physical Sciences Research Council (EPSRC) through the University of Warwick Innovative Manufacturing Research Centre. The authors also acknowledge the support of Ford Europe and the Lanner Group with respect to this research.


REFERENCES


AUTHOR BIOGRAPHIES
STEWART ROBINSON is Professor of Operational Research at Warwick Business School. He holds a BSc and PhD in Management Science from Lancaster University. Previously employed in simulation consultancy, he supported the use of simulation in companies throughout Europe and the rest of the world. He is author/co-author of three books on simulation. His research focuses on the practice of simulation model development and use. Key areas of interest are conceptual modelling, model validation and output analysis. Stewart is also carrying out research into understanding and modelling human interaction with systems. Home page <www.btinternet.com/~stewart.robinson1/SR.HTM>

ERNIE LEE received a BAcc from the Nanyang Technological University, Singapore. He has also completed an MSc in Operational Research from the Lancaster University Management School. At present, he is studying for his PhD at the Warwick Business School.

JOHN S. EDWARDS is Professor of Operational Research and Systems at the Aston Business School in the UK. He holds a BA and MA in Mathematics, as well as a PhD in Operational Research (Manpower Planning) from the University of Cambridge. His interest lies in understanding how people use models and systems to help them do things. For many years, he has studied the development and uses of knowledge based systems and decision support systems in management and administration, especially the integration of knowledge based systems, decision support systems and conventional information systems.