PACIS 1997 Proceedings

Pacific Asia Conference on Information Systems (PACIS)

December 1997

The "Sloppy" Modelling Of An Information Security Risk Model

Elizabeth Gurrie
Sunshine Coast University

Alison Anderson
Queensland University of Technology

Joachim Diederich
Queensland University of Technology

The "Sloppy" Modelling Of An Information Security Risk Model

Elizabeth Gurrie[1], Alison Anderson[2] and Joachim Diederich[3]

[1]Sunshine Coast University
Stringybark Rd and Sippy Downs Drive, Sippy Downs Queensland
Locked Bag No 4, Maroochydore South, QLD 4558
Phone: +61.7.5430 1217
E-mail: EGurrie@scuc.edu.au

[2]Information Security Research Centre
[3]Neurocomputing Research Centre
Faculty of Information Technology, Queensland University of Technology

Executive Summary

Object model development can be tedious and error-prone in contexts where the problem to be modelled is not well formulated, and particularly so in domains where concepts essential to its formulation may never have been previously defined. An object-based solution has many well-known benefits, so the choice *a priori* of an object-oriented implementation is not unusual. The modeller, however, must then capture objects, attributes, relationships etc. so as to exploit object-oriented benefits while preserving the "natural" model behaviours expected by the client. Incident threat analysis and prediction of computer security risk is a typical problem where the objects are difficult to identify and structure, because there is often no external consensus to model. One possible approach is "sloppy" modelling, which allows the modeller to postpone as late as possible decisions about object congruence. Its supporting methodology must therefore use mechanisms which promote convergence towards consensus, while exploiting multilevel user experience in the domain.

This paper presents a hybrid sloppy modelling approach to reducing the conceptualisation and structuring overhead in domains of this type. KARDS (Knowledge Acquisition for a Risk Data Store) was developed for and tested in a larger project for modelling organisational computer security risk. It represents an original solution which exploits hybridity to build stakeholders' individual perceptions of the problem into a unified repository which supports and interfaces with the object-oriented risk model. KARDS facilitates participation by multiple domain experts, speeds domain characterisation, and integrates structured and unstructured knowledge sources. KARDS is presented in the context of a typical threat incident analysis in the domain.

Introduction And Background To The Problem

The concept of "sloppy" modelling [Morik, 1988] acknowledges a difficulty inherent in knowledge capture, ie that a complete domain theory is presupposed. In accepting the need for a sloppy approach we are recognising that a well organised model is derived through process. Knowledge modelling is frequently directed by a pre-selected choice of knowledge representation. The choice of an object-oriented solution is often made without much insight into the problem domain, being perhaps influenced more by decisions about assumed benefits of the implementation vehicle, or projected context of use. It can certainly be argued that an advantage of the object model is its inherent genericity. When an object-oriented approach is adopted, it is necessary to supply not only an object-modelling interface but also a well-founded methodology for integrating heterogeneous knowledge sources[1]. This methodology must enable the modelling interface to resolve apparent contradictions, identify and tolerate temporary knowledge gaps, and consolidate unstructured domain concepts with any available structured domain knowledge. Our methodology acknowledges Morik's *sloppy* modelling principle and supports the evolution of domain theory from heterogeneous knowledge sources.

[1] The term *knowledge source* is used in this paper to refer to the origin of the knowledge acquired. This is in contrast to the use of the term by Breuker *et al* (1987) where knowledge-source refers to a domain relation that is the basis for some type of inference.

The domain of computer security risk assessment is one which lends itself to a *sloppy* approach. Because every organisation has its own objectives and values, every context is different, so modelling computer security risk becomes the task of modelling the interaction between the organisation and its information technology resources. Assessing a possible security threat and estimating its impact usually involves time-consuming and inexact scenario analysis by "experts" from non-intersecting fields of expertise (auditing versus data communications, for example). What is constant about the domain is that one expert's risked "asset" may be another expert's "threat" or even "countermeasure", and that risk may be multifaceted, quantitative or qualitative, dynamic, transient or static. The work described in this paper was carried out in conjunction with a larger project [Anderson *et al.*, 1993] based on the presumption that object-oriented representation is, for the general reasons cited above, ideal for the eventual formal risk model, as long as the "risk-related objects" themselves can be adequately captured along with as much as possible of their complex interrelationships. KARDS (Knowledge Acquisition for a Risk Data Store) specifically addresses the task of adequately capturing the knowledge associated with the impact of any breach to information system security. These impacts may be expressed in terms of confidentiality, integrity or availability and may be associated with system processes and transactions; or data assets and their complicated interrelationships. The impact of a security breach can manifest itself within any number of organisational functions (departments).

This paper begins by discussing the origins of the knowledge acquisition principles used in the development of KARDS, an interface to the object-oriented risk model, followed by an outline of the threat incident to be analysed, the *payday incident*. The major features of KARDS are presented, followed by a discussion of the results of applying KARDS to the *payday incident* threat. The final section evinces the success of KARDS and outlines a course for further work.

Knowledge Acquisition Principles Used In Kards

This section presents the fundamental modelling theory used in KARDS. The tool encapsulates the features of knowledge modelling which have been recognised as necessary in order to develop an accurate model of unstructured knowledge. The "nature" of the problem domain should act as a focus in deriving a model of requisite knowledge. Alexander *et al.* (1986) first suggested ontological analysis, and its importance has since been recognised by many researchers, notably Wielinga *et al.* (1993) and their ESPRIT project KADS (Knowledge Acquisition and Design Structure) and the subsequent CommonKADS project [Wielinga & Breuker, 1986; Hayward, 1987; Hayward *et al.*, 1988; Wielinga *et al.*, 1993; Schreiber & Wielinga, 1993; Schreiber *et al.*, 1994]. Knowledge-level analysis, conceptual modelling, expert participation and domain characterisation have provided the focus in our design of KARDS. However, this has been developed in the context of a *sloppy* approach. That is, that a domain model will evolve from the sloppy model constructed by the user [Morik, 1988].

With a view to satisfying the requirements of knowledge acquisition for the risk domain, KARDS aims to support the flexibility of an object-oriented risk model by applying knowledge acquisition techniques which have been demonstrated to overcome the type of problems seen in risk analysis [Gurrie *et al.*, 1995c].

The tool supports progression through a series of levels. This terminology is used to reinforce the fact that knowledge acquisition is an iterative and constructive task. It is most likely that each aspect of the acquisition process will be "revisited": this is the nature of knowledge. Acquisition of one chunk of knowledge will often indicate the need for another.

Each level of KARDS generates some aspect of knowledge which is then used at the next level of acquisition. The *Identification* level will specify the knowledge sources which should be accessed. The *Preparation* level will use those sources to prepare mediating knowledge [Nwana *et al.*, 1994], that knowledge which can be used by experts to convert their mental models to conceptual models. The *Elicitation* level uses the mediating knowledge to formulate conceptual models which are in turn stored in an interim knowledge base. Finally, the *Analysis* level will study the interim knowledge to generate encoded knowledge for an object-oriented risk model. The contribution made by our methodogy and its implementation, KARDS, is an original and unique application of knowledge acquisition theory to the information security domain.

At this point it is appropriate to confirm the understanding of knowledge acquisition assumed in this paper. Knowledge acquisition involves domain characterisation and the elicitation of knowledge structures. Knowledge elicitation and knowledge analysis differ in that elicitation results in informal knowledge structures, whilst analysis produces formal structures suitable for encoding [Shaw & Woodward, 1993]. As this tool is specific to the information security risk domain, domain characterisation is inherent. The levelled approach recognises the difference between elicitation and analysis; hence, elicitation methods will not be limited by representation issues. A knowledge-level analysis has revealed that heuristic classification [Clancey, 1985] •and the role-limiting method *acquire-and-present* [McDermott, 1988] will assist in eliciting the knowledge needed for a risk model.

Research from the computer security community has identified a number of players, or stakeholders, in risk analysis [Kowalski, 1990; Baskerville, 1991; Wahlgren & Carroll, 1992; Anderson, 1991; Ymström, 1992]. A focus of the KARDS approach is to ensure the active participation of each of these stakeholders. Our methodology assists by explicitly recognising that a variety of knowledge sources must be accessible to the stakeholders.

The Methodology

Figure 1 illustrates our methodology and its *sloppy* approach to the acquisition of information security risk knowledge.

Identification Level: At the identification level the sources of risk knowledge are distinguished. Identification is achieved by first considering those organisational or functional issues which may have some impact on information security. These issues may be ethical, legal or political, administrative, managerial, and operational [Kowalski, 1990]. This issue focus guides the identification of functional operations and stakeholders in risk analysis. As knowledge sources are identified, new knowledge about additional sources is acquired, an approach which supports "knowledge-guided elicitation".

Preparation Level: At the preparation level experts are assisted in verifying mental models through the processes of domain characterisation and expert training. This active role enables the validation of knowledge on an ongoing process and reduces the knowledge verification requirements of an object-oriented risk model. An outcome of this process will be mediating knowledge which assists in communication and conversion of mental models to conceptual models. Domain characterisation involves the development of a domain vocabulary, glossary, structure, purpose, and history [Nwana *et al.*, 1994]. These constructs can be elicited from human experts via some interview process such as repertory grid, or through text analysis of documents and transcripts. The automated analysis of database schemas can also assist in acquiring domain terminology.

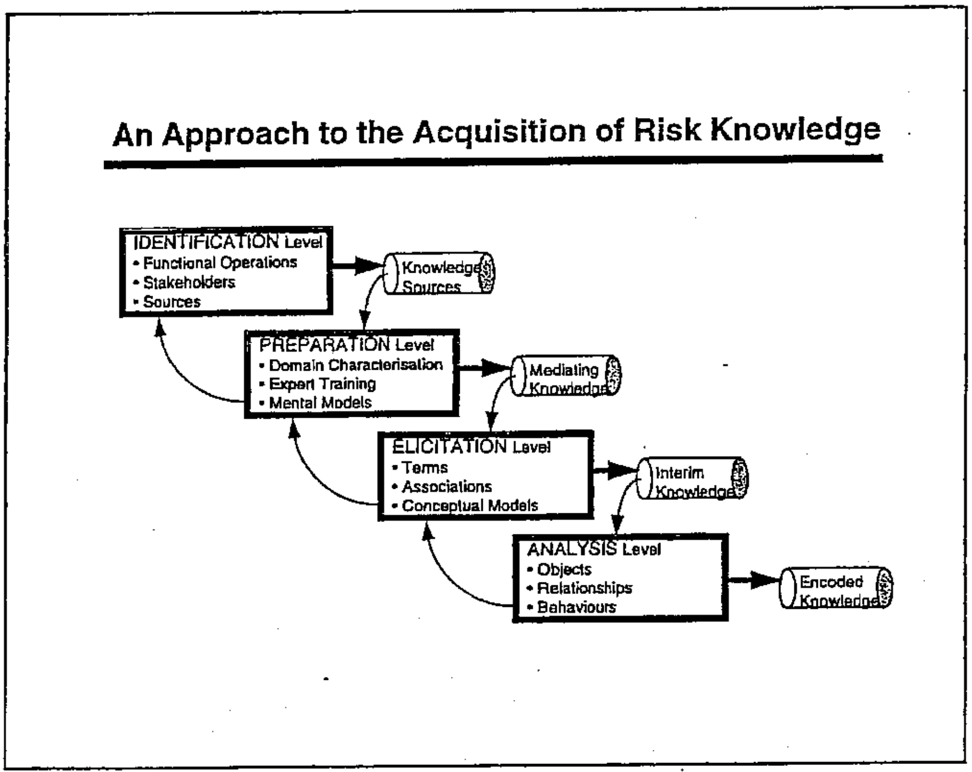


Figure 1 - A *Sloppy* approach to the acquisition of Risk Knowledge

If a variety of knowledge elicitation techniques is to be employed, domain experts need to be trained in the use of these. This will ensure an appropriate environment is established for the effective elicitation of knowledge structures and will help to overcome the effects of knowledge filtering, a problem arising from a knowledge engineer using their own interpretation of the expert's mental model [Anderson, 1993; Sandahl, 1994]. Another strategy for minimising knowledge filtering is the establishment of mediating knowledge [Nwana *et al.*, 1994]. Mediating knowledge supports communication between the domain expert and the knowledge engineer. Elicitation techniques which support the preparation of mediating knowledge include repertory grid and concept maps

Elicitation Level: The mediating knowledge generated at the preparation level can now be used to support the ongoing elicitation of domain knowledge. At this level experts use a variety of elicitation techniques to identify knowledge terms and their associations. A variety of knowledge sources should be accessible; hence, automated elicitation techniques such as database schema analysis, learning algorithms and text analysis were candidate techniques for implementation in our supporting acquisition tool.

The elicitation process results in the structuring of a number of conceptual models, from different domain experts. These conceptual models are integrated into an interim knowledge base [Diederich *et al.*, 1987] ready for analysis.

Analysis Level: Delivery of the target system is supported by the integration of interim knowledge into the object-oriented risk model. At the analysis level, interim knowledge is examined and encoded according to the object paradigm.

The Implementation Of Kards

KARDS was designed to assist in identifying knowledge about elements of the risk scenario. This knowledge may be the *risked-objects* and how the objects relate to each other, or less structured knowledge such as who manipulates the *risked-objects* or has an interest in them, the focus of this interest and the context in which the objects are at risk. KARDS does not attempt to assign a quantitative assessment of risk. If necessary, this will be undertaken by the parent system, the RDS [Anderson *et al.*, 1993]. However, our methodology does aim to acquire knowledge pertaining to any qualitative assessment of risk.

- At the *Identification* level, KARDS helps to distinguish available sources of knowledge about risked-objects. This level aims to assist users in identifying those personnel expert in some aspect of the system for which the risk analysis is to be undertaken and the functional operations which may be involved. During this *Identification* level, users may identify and access a variety of knowledge sources eg database schemas, existing interim knowledge and previous interview notes.
- At the *Preparation* level, KARDS assists in establishing the domain characterisation. Domain characterisation involves the development of a domain vocabulary and structure. Domain constructs are elicited from human experts via semi-automated interview. The automated analysis of database schemas also contributes to the development of the domain terminology. A thesaurus tool is available for further development of the domain terminology.
- At the *Elicitation* level, KARDS uses a variety of techniques for building associations between the conceptual models of information assets at risk as identified by the users. Elicitation is used to reconcile unstructured concepts and concept associations with codified versions of the same domain entity, eg the relationships implicit in a database schema.
- KARDS supports the *Analysis* level through the automatic derivation of additional relationships from the interim knowledge base. This is achieved by using existing interim knowledge to create training patterns for a neural network.

The knowledge acquired from all users is fully integrated by KARDS and can then be encoded in a form suitable for use by an object-oriented risk model.

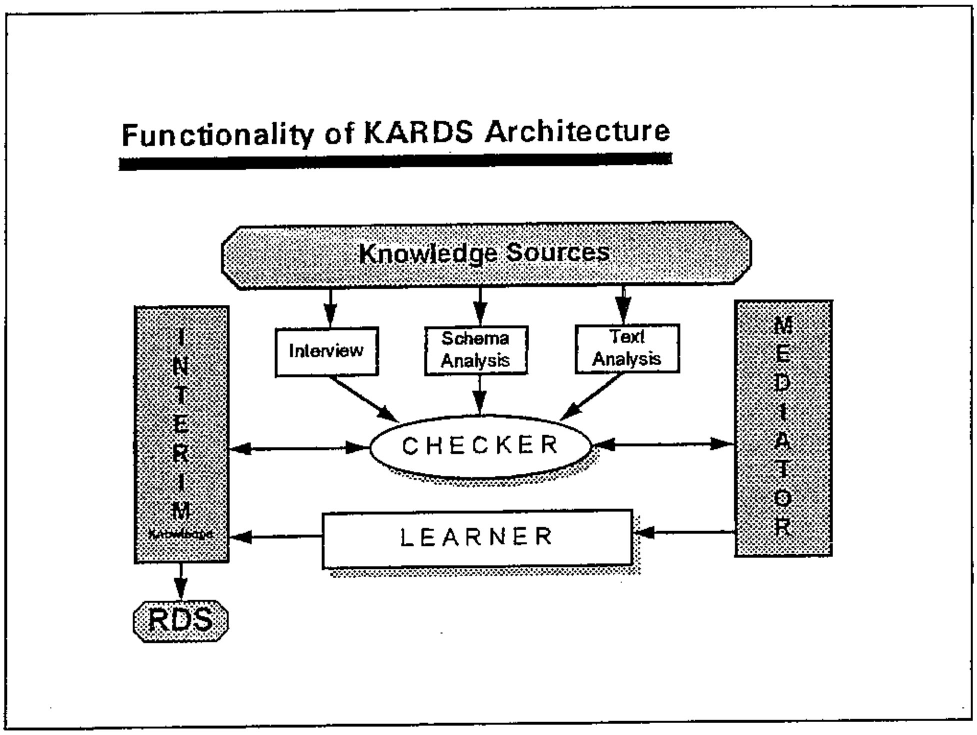


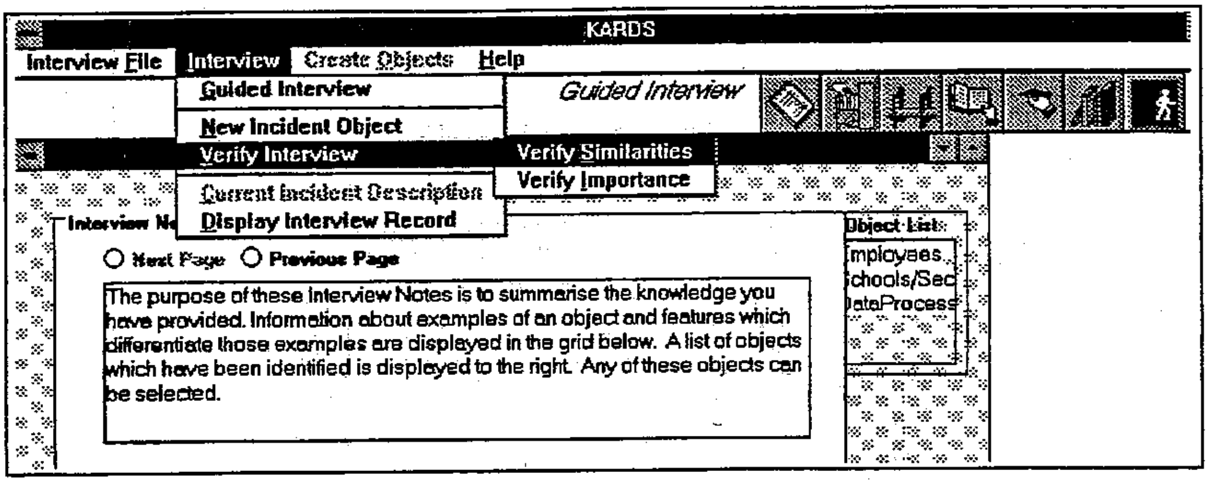
Figure 2 - The Architecture of KARDS

The Architecture Of Kards
The KARDS architecture, an extension to KRITON [Diederich *et al.*, 1987] and represented in figure 2, assumes access to a variety of acquisition methods. Both semi-automated and automated techniques are used to maximise the functionality of KARDS [Gurrie *et al.*, 1995b] and support the methodology described in this paper.

As knowledge components are acquired, the *Mediator* assists the experts in developing a conceptual model of their approach to, and interest in, risk information. As knowledge is acquired and mediated, it is verified by the *Checker*. Mediation and checking allow the development of an *Interim Knowledge Base*, an integration of the different conceptual models offered by the various stakeholders contributing to the knowledge acquisition task. The *Learner* supports the principle of knowledge-guided elicitation. Pre-elicited knowledge serves as examples which can be used by the *Learner* to derive new knowledge. The interim knowledge is encoded in a form suitable for an object-oriented risk model, providing an environment which supports the evolution of a domain theory.

The *Mediator*: The initial role of the *Mediator* is to instigate an interview. An event, or incident, is considered to be a specific instance of a threat. This is used to provide a focus for the security expert during the interview. Icons, concept graphs, grids and text are other examples of presentation techniques used by the *Mediator* to assist the users in communicating their mental model.

The *Checker*: The time to undertake knowledge consistency checking is during the knowledge acquisition process [Geissman & Schultz, 1988]. A knowledge acquisition tool should provide the opportunity to *confirm the model's capacity to describe correctly and accurately the problem situation* - and *the quality and applicability of the solutions and recommendations* [O'Keefe & O'Leary, 1993]. The *Checker* supports this principle at each level of knowledge acquisition in addition to providing the functionality of the WATCHER as implemented in KRITON [Diederich *et al.*, 1987].



The purpose of these Interview Notes is to summarise the knowledge you have provided. Information about examples of an object and features which differentiate those examples are displayed in the grid below. A list of objects which have been identified is displayed to the right. Any of these objects can be selected.

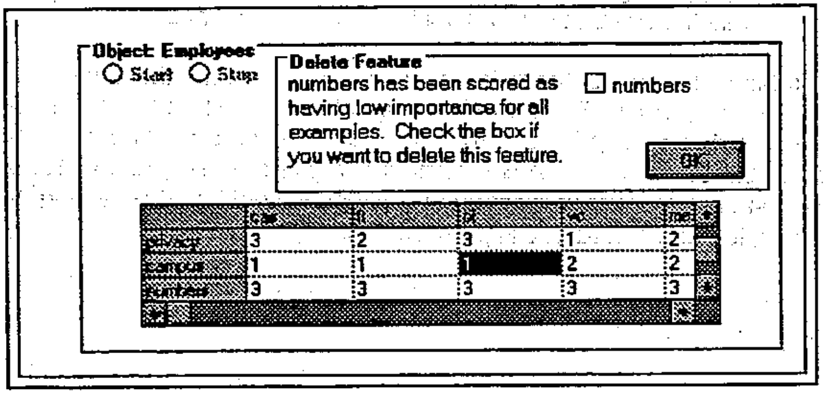
Figure 3 - The Checker *Verifies Interview Knowledge*

An initial task for the *Checker* is to verify the knowledge acquired via interview (see Figure 3). It will check for apparent redundancies (see Figure 4) or incompleteness in the interview transcript. In addition, the *Checker* provides the facility to verify concept associations. These are either heuristically or manually acquired and are verified by the user to ensure the accurate acquisition of the organisation's security model.

The *Interim Knowledge Base*: The knowledge acquired by any of the acquisition methods used in KARDS is stored in an interim knowledge base. This interim knowledge represents an integration of all conceptual models. It is continually updated and restructured as new knowledge is acquired.

The *Learner*: In many application areas, domain knowledge is not only available in the form of rules or facts, as provided by an expert or the analysis of relevant documents, but also in the form of data or cases which represent prior experiences in similar situations. It is natural to combine knowledge-driven and example-based acquisition methods. Neural networks are candidates for "learning by example" acquisition techniques. KARDS implements a neural network to support data-analysis and feature extraction [Gurrie *at al.*, 1995b], ie KARDS discovers previously unknown relations and attributes in data sets. These relations can enrich the object knowledge acquired by other KARDS knowledge acquisition techniques.

An additional role of the *Learner* is to acquire and maintain the domain vocabulary. Each elicitation method results in the acquisition of new terminology. This is recorded, maintained and linked to associated terms.



numbers has been scored as having low importance for all examples. Check the box if you want to delete this feature.

	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	3	2	3	1	2
[illegible]	1	1	1	2	2
numbers	3	3	3	3	3

Figure 4 - The *Checker* Identifies Irrelevant Information

Overview Of The "*payday incident*" Threat

The *payday incident* has been selected to demonstrate the effectiveness of the KARDS architecture. Briefly, the case-study poses the task of identifying those information assets which may be involved, or impacted on, by some security incident on pay day in a university environment. A number of organisational functions are likely to be involved, eg human resource section, finance department, payroll section, computer services section, and general administration. Multiple users, policies and risk perceptions are therefore present.

The *payday incident* may involve the failure of component software; failure to process pay due to hardwáre faults; failure to process electronic transfer of pay due to data communications failure; failure to process pay due to industrial action by staff; inability to process pay due to unauthorised or accidental destruction of database; or some other incident which prevents accurate processing of salaries. Hence, multiple versions of the threat incident itself are present as well. The task is to identify what assets are potentially risked and the impact of such an incident in legal, political, ethical, administrative, managerial or operational terms. This unstructured information should be ascertained by eliciting relevant knowledge from appropriate organisational sources. As a minimum, any asset identification process should identify the systems, particularly the software and data, which may be at risk in the *payday incident,* eg the computer systems responsible for processing pays; the system components, both hardware and software; the database resources used for pay detail records.

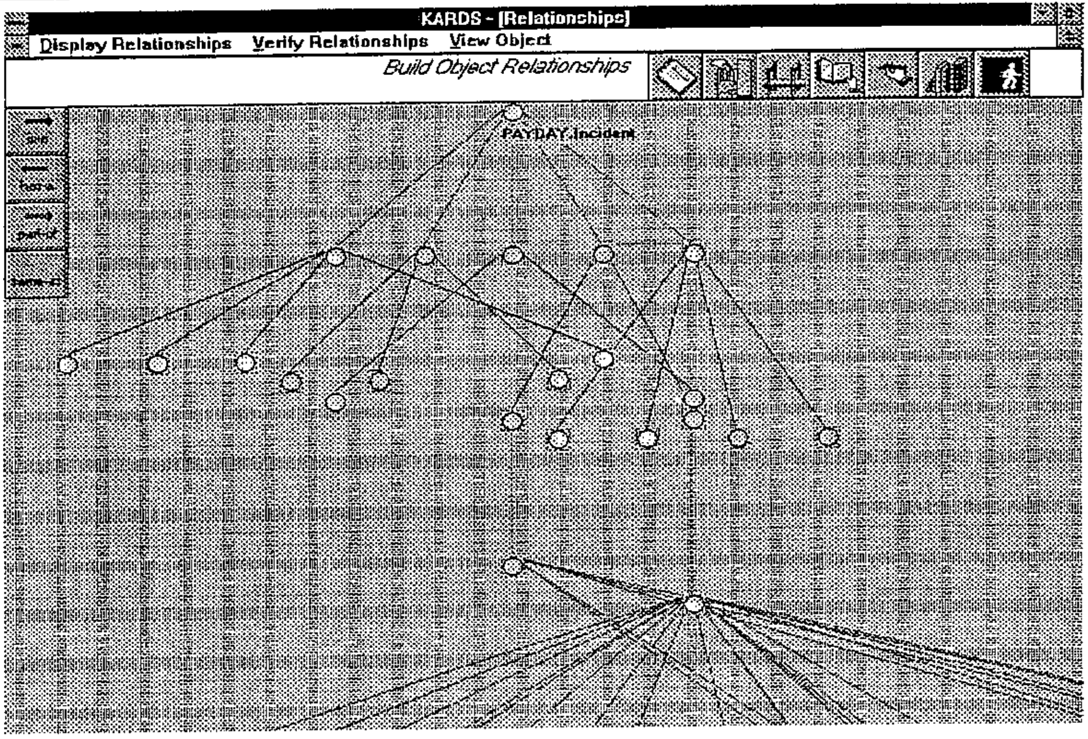
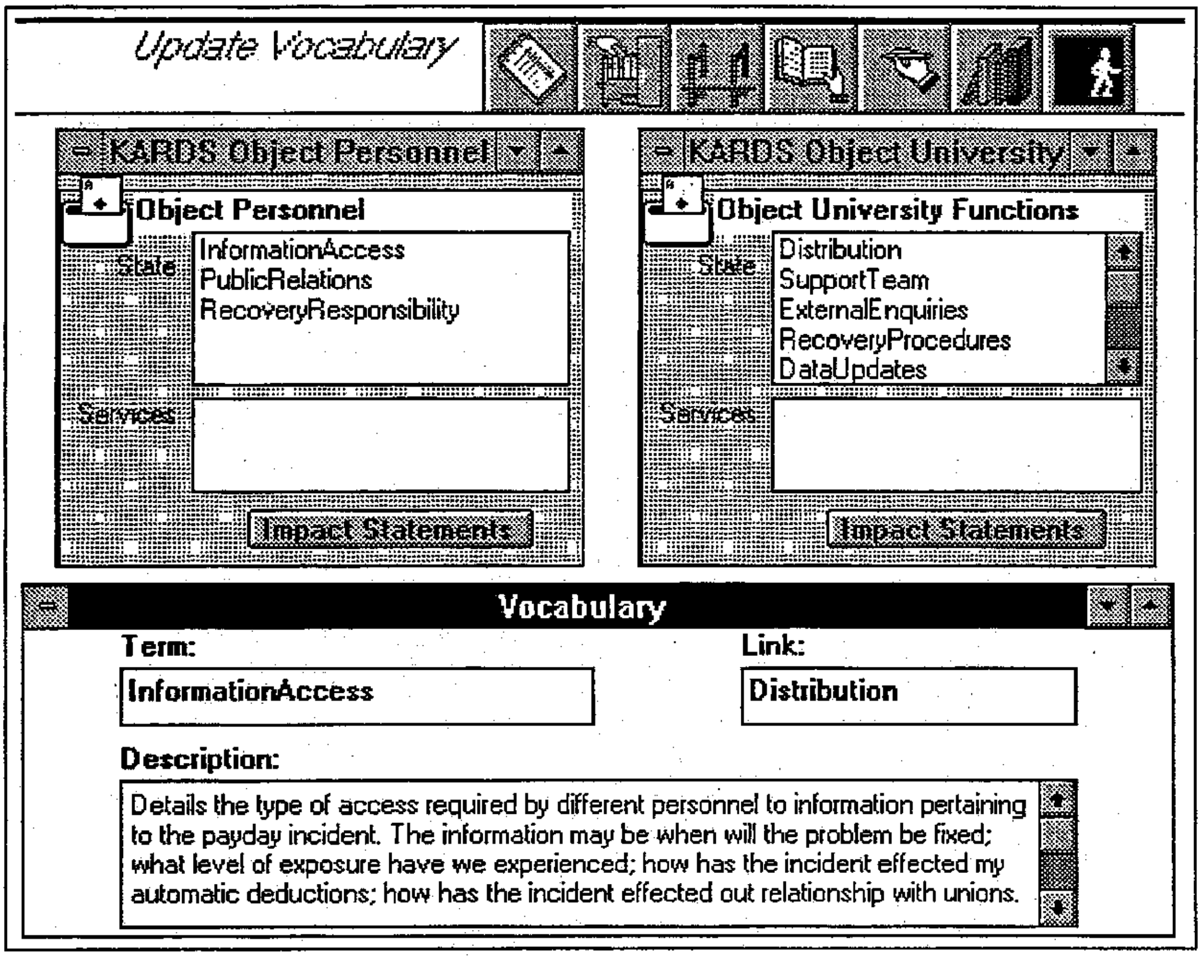


Figure 5 Concept model of payday incident *as generated by KARDS during knowledge acquisition process*

Application of KARDS to the *payday incident* Analysis

Figure 5 shows the concepts, and their relationships, which have been identified through a series of knowledge capture sessions using KARDS. The KARDS *Mediator* allows any of these concepts to be selected and expanded by the user if required. This conceptual model integrates knowledge acquired from a variety of sources and demonstrates the contribution of each level of our methodology. This section outlines the use of our sloppy methodology in evolving a concept model.

Identification: The objects generated by KARDS reflect different perceptions of the *payday incident*, and the ability of one user to elaborate on the contribution of another. For example, the interview process began with the Human Resources (HR) Manager. The HR manager has identified that several personnel could be involved in a security incident on payday.



Update Vocabulary

KARDS Object Personnel

Object Personnel

State

InformationAccess
PublicRelations
RecoveryResponsibility

Services

Impact Statements

KARDS Object University

Object University Functions

State

Distribution
SupportTeam
ExternalEnquiries
RecoveryProcedures
DataUpdates

Services

Impact Statements

Vocabulary

Term: InformationAccess

Link: Distribution

Description:

Details the type of access required by different personnel to information pertaining to the payday incident. The information may be when will the problem be fixed; what level of exposure have we experienced; how has the incident effected my automatic deductions; how has the incident effected out relationship with unions.

Figure 6 - Developing payroll incident domain characterisation via the KARDS thesaurus

Preparation: Identification has introduced a number of terms which need to be defined in order to ensure accurate representation of the knowledge. For example, Figure 6 demonstrates domain characterisation by showing how the term *InformationAccess* is defined by the HR Manager. Another user has subsequently created a link from *InformationAccess* to the term *Distribution*, indicating they are related.

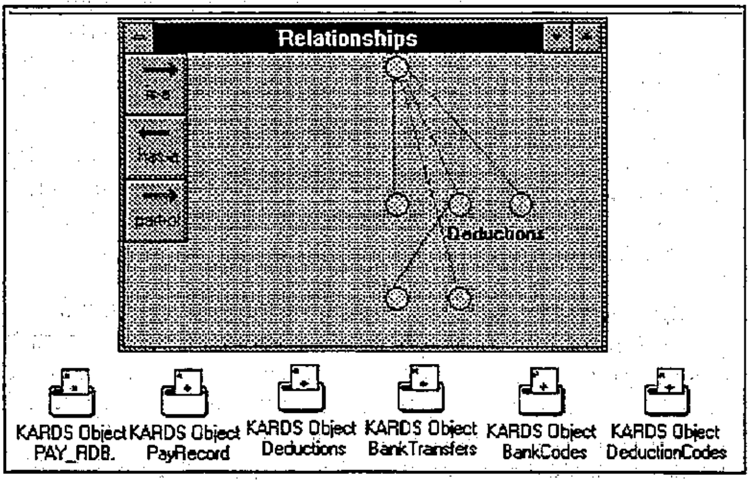
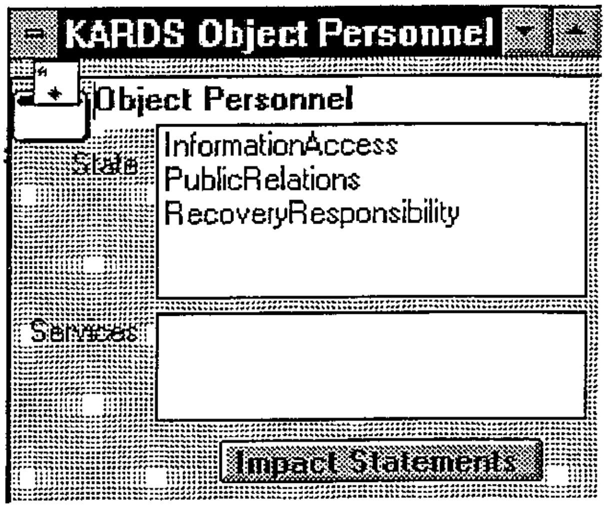


Figure 7 - KARDS derived relationships from analysis of payroll database schema

Elicitation: Figure 7 displays the objects automatically elicited by KARDS' analysis of the payroll database schema. In fact, during "Elicitation", KARDS has heuristically analysed a given set of database schemae to produce a suggested relationship hierarchy. This can be corrected or manipulated by the user directly, or a semi-automated "laddering" process [Boose & Bradshaw 1988a, Corbridge *et al.*, 1994] can be used to distinguish the type of relationship e.g. *is-a, has-a, part-of.* As subsequent users interact with KARDS, this relationship hierarchy can be integrated with any new knowledge acquired.

Analysis: Ambiguities, lacunae and undiscovered relationships remaining after elicitation are reconciled in KARDS through its implementation of a neural network simulator in its *Learner* [Gurrie *et al*, 1995a]. In the *payday incident*, the *Learner* identified that objects *personnel* and *employees* were identical, where the object *personnel* was interactively captured via interview at the *Identification* level, and *employees* was automatically derived from heuristic schema analysis at the *Elicitation* level. The nature of the involvement of these personnel has been defined by the user as being associated with their ability to access information relating to the incident, their responsibility for public relations associated with the incident and their responsibility for ensuring recovery from the incident. The KARDS *Mediator* has presented this interim knowledge as depicted in Figure 8. The *Learner* subsequently refines this knowledge.



InformationAccess
PublicRelations
RecoveryResponsibility

Figure 8 - Interim knowledge elicited from the Human Resource Manager

Additional concepts acquired include knowledge associated with the impact on bodies external to the university e.g. banks, unions, media and insurance companies. Further interviewing may result in defining the involvement of university functions such as human resources, finance department, payroll section, faculties or divisions, schools or sections and the information desk. Following is a summary of the type of knowledge which was acquired by KARDS in relation to the *payday incident*.

Information Assets: KARDS does not exclude any knowledge from the information assets identified by a user; the KARDS interface is unstructured so that any knowledge which the user considers relevant can easily be acquired. The *payday incident* scenario has demonstrated that KARDS is able to identify many different types of knowledge. Structured knowledge such as software and data assets can be identified, along with physical assets. The acquisition of unstructured knowledge is supported such as contextual knowledge about the interaction between organisational units and software and data assets they are using, the responsibilities personnel may have in association with data assets, and any interaction that may occur with external bodies and their expectations of the reliability of data assets.

Conceptual Model: KARDS is successful in maintaining the relationships between information assets which have been identified. Figure 5 presents an object relationship model generated by KARDS at one stage in the acquisition process for the *payday incident*. In order to indicate more of the detail in this model, Figure 9 illustrates the objects identified and the relationships between those objects. Figures 5 and 9 demonstrate the richness of the knowledge KARDS is able to maintain.

Impact Statements KARDS supports the acquisition of qualitative impact statements. These impact assessments are drawn from knowledge obtained through domain characterisation, ie the KARDS thesaurus, and direct user input via the passive business impact statements recording facility.

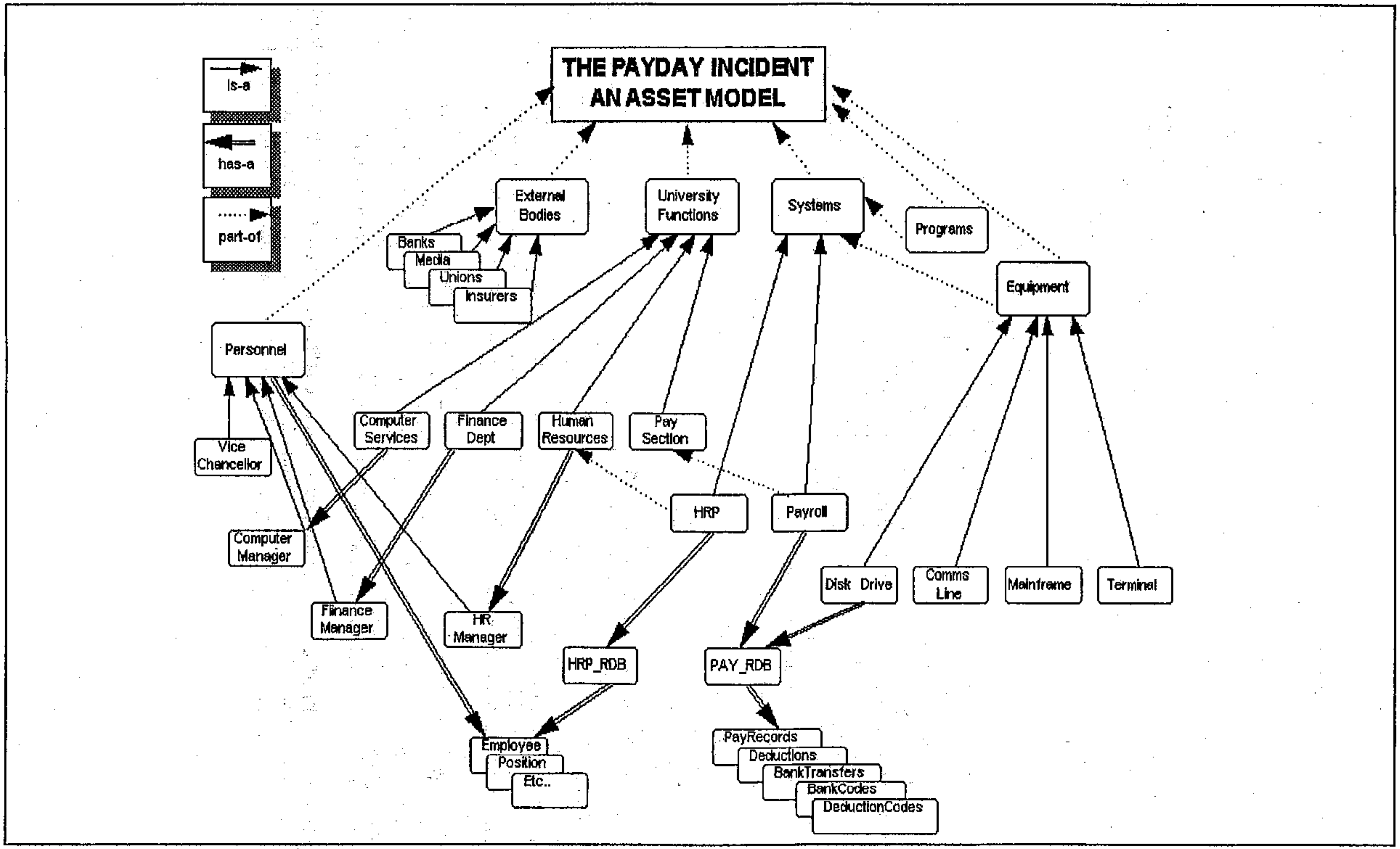


Figure 9 - The Asset Model generated by KARDS for the payday incident

Comments On Kards Output

KARDS' asset identification reflects an integration of different user perspectives of the *payday incident*. The assets identified include new and inferred knowledge acquired automatically, and indicate KARDS' ability to identify *difficult-to-capture* knowledge.

All knowledge acquired and retained by KARDS is used in the final output. This is managed through user control. Knowledge can be identified as redundant either automatically or manually, and the user decides whether redundant knowledge is deleted or modified, thus ensuring the final output is an accurate reflection of the stakeholders' mental models.

Qualitative assessments of the impact of some security event can be assigned at any level of the knowledge structure. This supports the architecture of the parent system, the RDS, and assists in minimising the information collection process since impact assessments will be applied to related objects by the RDS.

KARDS provides a supportive and interactive environment allowing users to reflect upon the contribution of other users. This assists in minimising duplication of information and maximising collaborative input. The methodology underlying the design of KARDS ensures active participation of experts, and a collective approach to identifying *objects-at-risk*. A variety of knowledge presentations are supported by the KARDS *Mediator*. This provides flexible feedback to users, supporting the exploration of their concepts and their contribution to the overall security model.

KARDS generates a rich knowledge structure which exposes and formalises the hierarchical relationships between *risked-objects* and their environment. This ensures interdependencies between *risked-objects* are reflected in the asset model generated. Importantly, KARDS forces no pre-determined structure on knowledge; rather, it relies on a process of refinement of the conceptual and informal knowledge in order to develop the structure.

Evaluation And Discussion

This section presents a summary of the results obtained in an initial evaluation of KARDS' performance in the *payday incident*. These qualitative comments provided by users expert in the domain suggest KARDS offers a feasible approach to solving the problem of modelling in unstructured domains such as computer security risk modelling.

The completeness of the knowledge acquired by KARDS was considered to be high. The KARDS tool assisted in the acquisition of knowledge relating to personnel, university functions, responsibilities, procedural knowledge, software, hardware and database assets. This knowledge was acquired from a variety of formal and informal sources and integrated into a single conceptual model.

The accuracy of the knowledge acquired was considered to be dependent on the expertise and individual perceptions of the users. However, KARDS assisted users in optimising the accuracy of their model by checking for any redundancy. Importantly, the KARDS features which support domain characterisation assisted in identifying any overlap in acquired knowledge. Such overlap was retained by the KARDS thesaurus and referenced by the *Learner*. Further, where contradictions in the knowledge were uncovered by the CHECKER, users were supported in resolving these contradictions.

Approximately half of the knowledge acquired through automated database schema analysis was not considered to be directly relevant to the *payday incident* by users. Hence, the LEARNER has the potential to generate some irrelevant knowledge, this has not been fully investigated as yet.

An important aspect of computer risk analysis is the identification of information asset relationships. KARDS generated a comprehensive hierarchy of asset relationships. It acquired knowledge about multiple relationships between *risked-objects* and supported the encoding of inheritance.

Automated schema analysis prevented any requirement to manually record database objects as information assets to be protected. Semi-automated interview negated the need to record an interview manually then enter acquired data. Heuristic relationship building minimised the need for user identification of relationships.

Users were particularly satisfied with the level of guidance offered within KARDS. The tool's architecture supported the acquisition of both structured and unstructured knowledge. Users expressed satisfaction with the flexible environment of the *Mediator* and its ability to provide alternative representations of user mental models. Of greater importance is KARDS' ability to support the integration of different user concepts. Users' conceptual models were fully integrated. Further, the knowledge acquired automatically was integrated with user concept models to form an interim knowledge base.

This evaluation indicates the KARDS architecture offers a novel solution to the problem of acquiring knowledge associated with ill-defined *risked-objects.* The KARDS architecture offers a natural and flexible environment, identifying relevant, non-redundant knowledge. In the context of a given typical problem, KARDS correctly discovered concepts central to the scenario, revealed both known and unknown relationships among them, disposed of redundancies, and successfully generated an interim object base suitable for updating the formal security risk model.

Risk analysis has traditionally been considered the bottleneck in security risk modelling; it has been demonstrated that KARDS ability to *directly* access a variety of knowledge sources helps to minimise this problem. Further, the hybrid nature of the knowledge acquisition techniques available in KARDS provide an original approach to the acquisition of different types of knowledge from multiple sources in the information security domain.

Future Work

There are two main areas of further work revealed by the experience in this limited application of KARDS. Firstly, the methodology and its implementation require additional testing with alternative threat incidents, and the same threats in a different organisational context. Further, the interaction between the KARDS-produced interim object base, and the permanent object-oriented risk model it updates is clearly under-exploited as a knowledge source in its own right.

As implemented, update proceeds from KARDS to the structured risk model, in one direction. However, the structured object model of the latter could also be used within KARDS as a knowledge source which is known already to contain "clean" contextualised risk knowledge. Access to a base of pre-existing formally defined *risked-objects* could be expected to provide incremental efficiency gains in subsequent knowledge capture.

- It could also be expected that the experimental behaviour of KARDS itself reveals useful aspects of the domain that can be not just incorporated in an improved KARDS, but also result in a more streamlined approach to "sloppy" modelling itself. KARDS satisfactorily carries out concept convergence - eliminating redundancy, revealing gaps and exposing apparently irrelevant concepts - but a more fundamental question is whether aspects of its convergence behaviour reveal special domain characteristics that could be used to improve the "risk-specificity" of the methodology, for example:
- the current mode of redundancy detection and resolution may be an indication of how well the threat incident acts as a prototypical threat scenario, ie, a high redundancy might indicate a good focus in the sense of extracting as much risk-relevant information out of a single case. This line of enquiry might reveal a likely prototype "threat incident" framework for use in early stages of the methodology;
- similarly, the rate of convergence: does this represent a measure of how well the problem is understood, or the reliability of the "expertise" being tapped? In either case, it might be a measure of the relative semantic contribution of the selected incident, and might therefore be used to evaluate the incident's effectiveness in knowledge elicitation;
- knowledge currently identified as "irrelevant" is revealed in the problem analysis, but discarded. Is it really irrelevant, or does it indicate a contingent threat that should be analysed in conjunction with the current one: could a fully-automated methodology learn to direct its own line of enquiry, and optimise its own levels?

Summary

In this experiment we have investigated the confluence of information systems theory and practice. Well-established techniques from knowledge acquisition theory have been applied to the problem of object identification and capture in a complex and poorly structured domain with multiple knowledge sources. The *payday incident* case study demonstrates both the feasibility of a hybrid interface and the benefits of such an approach in attacking the "knowledge capture bottleneck" intrinsic to the problem domain. KARDS successfully generates an interim object base tailored for a specific context. These objects are sufficiently detailed and in a suitable form for updating the generic object-oriented risk model which forms a permanent, dynamic representation of organisational information security risk.

References

Alexander, J, Freiling, M., Shulman, S, Staley, J., Rehfuss, S. & Messick, S. (1986) "Knowledge Level Engineering: Ontological Analysis" *Proceedings of AAAI-86* pp963-967

Anderson, A. (1991) "Comparing Risk Analysis Methodologies" in **Lindsay, D.T. & Price, W.L. (Ed)** *Proceedings of IFIP TC11 Seventh International Conference on Information Security: Creating Confidence in Information Processing, IFIP Sec '91 Brighton, UK, 15-17 May 1991* Elsevier Science Publishers, North-Holland

Anderson, A., Longley, D. & Tickle, A.B. (1993) "The Risk Data Repository: a novel approach to security risk modelling" *Proceedings IFIP Sec '93*, Deerhurst, Canada.

Anderson, J.R. (1993) "Development of Expertise" in **Buchanan, Bruce & Wilkins, David (Eds)** *Readings in knowledge acquisition and learning: Automating the Construction and Improvement of Expert Systems*, Morgan Kaufmann Publishers, San Mateo, California.

Baskerville, R. (1991) "Risk Analysis as a Source of Professional Knowledge" *Computers & Security* v10 n8 pp749-764.

Boose, J.H. & Bradshaw, J.M. (1988a) "Expertise transfer and complex problems: Using AQUINAS as a knowledge-acquisition workbench for knowledge-based systems" in **Boose, J. & Gaines, B.**, *Knowledge Acquisition Tools for Expert Systems: Knowledge-based Systems Vol 2* Academic Press Ltd.

Clancey, W.J. (1985) "Heuristic Classification" *Artificial Intelligence* v27 pp289-350.

Corbridge, C., Rugg, G., Major, N.P., Shadbolt, N.R. & Burton, A.M. (1994) "Laddering: technique and tool use in knowledge acquisition" *Knowledge Acquisition* v6 pp315-341

Diederich, J., Ruhman, I. & May, M. (1987) "KRITON: a knowledge-acquisition tool for expert systems" *International Journal of Man-Machine Studies* n26 pp29-40.

Geissman, James & Schultz, Roger (1988) "Verification & Validation of Expert Systems" *AI Expert* Feb pp26-33.

Gurrie, E., Diederich, J., Tickle, A. & Anderson, A. (1995a) "Knowledge acquisition and the re-use of security risk knowledge" *Proceedings of The Knowledge Engineering Forum - Concepts and Architectures for Reuse March 30-31*, 1995, GMD Sankt Augustin, Germany, pp44-50.

Gurrie, E., Diederich, J., Tickle, A. & Anderson, A. (1995b) "KARDS: Hybrid knowledge acquisition for a security risk model" *Proceedings of Eighth International Conference on Industrial & Engineering Applications of Artificial Intelligence and Expert Systems (IEA/AIE - 95) June 5-9, 1995* Melbourne, Australia, pp501-506.

Gurrie, E., Tickle, A. & Anderson, A. (1995c) "A case for knowledge acquisition in risk analysis" *Technical Report #1/95, Faculty of Information Technology, Queensland University of Technology* QUT, Brisbane, Australia.

Hayward, S.A. (1987) "Models for Knowledge Acquisition" in **Diederich, J. & Uthmann, T.** (Eds): *Knowledge Acquisition for Expert Systems* Olderboing Verlag.

Hayward, S.A., Wielinga, B. J. & Breuker, J.A. (1988) "Structured analysis of knowledge" in **Boose, J. & Gaines, B.**, *Knowledge Acquisition Tools for Expert Systems: Knowledge-based Systems Vol 2* Academic Press Ltd.

Kowalski, S. (1990) "Rapport till IEE International Orange Book Workshop: The Security by Consensus Model, USA maj 1990", *Working Paper: 90-12 System Integrity and Information Security research centre,* Department of Computer and Systems Science, Royal Institute of Technology, Stockholm, Sweden.

Linster, M. (1993) "Integrating conceptual and operational modelling: a case study" *Knowledge Acquisition* v5 pp143-171.

McDermott, J. (1988) "Preliminary steps toward a taxonomy of problem-solving methods" in **Marcus, S.** *Automating Knowledge Acquisition for Knowledge Based Systems* Kluwar Academic Publishers.v5 pp143-171.

Morik, K. (1988) "Acquiring Domain Models" in **Boose, J & Gaines, B. (Ed)** *Knowledge Acquisition Tools for Expert Systems: Knowledge-based Systems Vol 2* Academic Press Ltd.

Nwana, H., Bench-Capon, T.J.M., Paton, R.C. & Shave, M.J.R. (1994) "Domain-driven knowledge modelling for knowledge acquisition" *Knowledge Acquisition* v6 pp243-270.

O'Keefe, R. & O'Leary, D. (1993) "Expert System Verification and Validation: a survey and tutorial" *Artificial Intelligence Review* v7 pp3-42.

Sandahl, K. (1994) "Transferring knowledge from active expert to end-user environment" *Knowledge Acquisition* v6 pp1-22.

Schreiber, A. Th. & Wielinga, B. J. (1993) "Comparing KADS to conventional software engineering" In **A. Th. Schreiber, B. J. Wielinga, and J. A. Breuker, editors**, *KADS: A Principled Approach to Knowledge-Based System Development*, pp 151-165. Academic Press, London.

Schreiber, A. Th., Wielinga, B. J., Akkermans, J. M., Van de Velde, W., and de Hoog, R. (1994) "CommonKADS: A comprehensive methodology for KBS development" *IEEE Expert*, 9(6), December 1994.

Shaw, M.L.G. & Woodward, J.B. (1993) "Modeling expert knowledge" in **Buchanan, Bruce & Wilkins, David (Eds)** *Readings in knowledge acquisition and learning: Automating the Construction and Improvement of Expert Systems*, Morgan Kaufmann Publishers, San Mateo, California.

Wahlgren, G. & Carroll, J. (1992) "General Systems Theoretic Model of InfoSecMan" *Proceedings of Workshop on Information Security Management IFIP, Technical Committee 11, Working Group 11.1, Carlton Hotel Singapore, May 1992.*

Wielinga, B. J. & Breuker (1986) "Models of Expertise" *Proceedings, European Conference on Artificial Intelligence* pp306-318.

Wielinga, B.J., Schreiber, A.T. & Breuker, J.A. (1993) "KADS: A modelling approach to knowledge engineering" in **Buchanan, Bruce & Wilkins, David (Eds)** *Readings in knowledge acquisition and learning: Automating the Construction and Improvement of Expert Systems*, Morgan Kaufmann Publishers, San Mateo, California.

Yrnström, L. (1992) "Towards a Systemic-Holistic Approach to Academic Programs in the area of IT Security", *DSV Research Report Series ISSN 1101-8256 No. 92–026DSV*, Stockholm University, Dept. of Computer and Systems Sciences, Stockholm.