Why Database AND Expert Systems?

Trevor J.M. Bench-Capon
Department of Computer Science
The University of Liverpool
Liverpool,
England

Abstract

In this talk I discuss the reasons why it is important for the database and expert systems communities to come together to exchange ideas and techniques. I discuss the lessons that each community can learn from the other, and envisage a future in which the two styles of system will be integrated, adding value to both.

1. Introduction

The series of Databases and Expert Systems Applications conferences has been running throughout this decade, bringing together what are normally regarded as two separate communities. The coverage of the conference is broad, and every aspect of databases and expert systems has been represented, along with such related matters as information retrieval, hypertext and the like. Nearly all of the papers, however, can be clearly identified with one or other of the technologies. Put another way the DEXA conferences, have tended to represent the union of Databases and Expert Systems concerns rather than the intersection. If this is so, and Databases and Expert Systems can been seen as two rarely intersecting strands running through the conferences, we must ask what is the *rationale* of having a single conference to deal with these two disparate subjects. In this paper I shall argue that it *is* important that the two communities come together, from my perspective, which is that of an expert systems person. I shall begin by discussing what the database community can teach the expert systems community. I shall then offer some pointers to work in expert systems which has relevance for database. Finally I shall argue that the information systems of the future will require a close co-operation between the two styles of system.

2. The Relation Between Databases and Expert Systems

Databases, and expert systems, both fall under the broad category of information systems. The approach they take, however, is markedly different. The database approach is to store a large amount of facts extensionally. In order to do problem solving using information in the database, an application program (or a user) must supply intentional definitions to extract the information that is required. An expert system, in contrast, stores a number of intentional definitions, and the user (or application program) must supply specific information about a particular case to which these definitions can be applied. They can thus be seen as complementing one another: problem solving can be seen as applying problem solving expertise to a set of specific facts relating to a particular case: in both systems general expertise is applied to specific information, but whereas the database supplies the specific information, the expert system is supposed to supply the expertise.

As a result of this difference of approach databases can be characterised as holding *large* amounts of data which must be available for use in a *variety of tasks,* whereas expert systems usually hold only a *small* amount of data, and are dedicated to performing a *limited range of specific tasks*. The research concerns of the two communities flow from this: in database we are concerned with efficient storage and retrieval, task neutral modelling, robust systems, etc, whereas in expert systems we are involved in capturing problem solving expertise, reconciling conflicting information, representing judgement, etc.

Another contrast is that databases form the basis of the commercial data processing industry, whereas even today, after a decade and a half of development, expert

0-8186-8147-0/97 $10.00 © 1997 IEEE

systems are still largely experimental, and form only a small part of the "real world" computing scene.

3. What Expert Systems Can Learn From Database.

This last point - the lack of commercial take-up - has given rise to considerable concern amongst workers in the expert system field. When the technology was new the following kind of things were being said:

> "The UK Civil Service is the largest single user of conventional IT equipment and services in the UK ... The CCTA has a specific responsibility to research and then encourage the use of appropriate IT to assist in the administrative mechanisms of Government. KBS represents one such technology which CCTA has identified as being of particular benefit ... In terms of government administration, KBS may be the single most significant development to emerge since the computer itself, for it offers a means of streamlining and improving decision-making to an unprecedented degree." (Duffin 88, page 7).

KBS (Knowledge Based Systems) and expert systems may be taken as synonymous for our purposes. The CCTA is the UK Central Computer and Telecommunications Agency, and Duffin was a fairly senior figure in it: we therefore need to explain why his prophecy was not realised. As a result of reflection on this matter, a number of problems inhibiting the routine use of expert systems have been identified:

- Difficulties in acquiring knowledge: uncovering the problem solving expertise required to produce a significant system is a painstaking and resource intensive task. The investment required is difficult to justify unless clear advantages can be shown to result;
- Re-use of knowledge: because of the effort involved in acquiring knowledge about a domain, there is a strong motivation to use knowledge acquired to solve one problem in a given domain in other systems in that domain. However, because expert systems tend to be constructed with a quite specific task in mind, this tends to be rather difficult.
- Verification, Validation and Maintenance: if a system is to go into real, everyday, use these aspects are of the utmost importance. Unfortunately the verification and validation of expert systems is still an area of active research rather than one where there are well understood solutions, and the maintenance of expert systems is a problem which is as yet hardly addressed.

All of these matters have, of necessity, been thoroughly explored in relation to databases. In particular the very idea of databases requires that the information be capable of use by a variety of applications.

The key to the success of databases is the robust data model (schema) on which they are based. When designing a database this is seen as of crucial importance, and it provides the foundation for all use of the database. This contrasts with expert systems constructed on traditional lines, where the data model receives scant, if any, attention. In traditional knowledge engineering, knowledge, typically in the form of rules, is elicited from an expert. These rules reference data, and so a data model emerges from the rules, but it is very often incomplete, and always somewhat haphazard. For example, if the rule contains reference to sex being male, an attribute sex with a potential value of male is required. But the alternative values, if not referenced by a rule condition, will remain implicit. It may, however, be important to know whether the alternatives are female only, or either female or neuter. This can be particularly crucial if we want to make use of the negative information that sex is not male.

The importance on the data model is a crucial lesson for expert systems, and in recent years expert systems people have begun to realise this. Often today the realisation is expressed as argument for the importance of *ontologies*. An ontology for an expert system is perhaps best described as "an explicit specification of the conceptualisation of its domain" (Gruber 1995). If a system commits to an ontology it agrees "to use a vocabulary (i.e. ask queries and make assertions) in a way that is consistent (but not complete) with respect to the theory specified by an ontology" (Gruber 1995). Those familiar with database schemata will notice the strong resemblance, and the use made of it by application programs. The real importance of ontologies is that producing them forces the expert systems builder to place an initial focus on design and domain modelling, and results in a product which provides explicit documentation for the expert system and the assumptions which underpin it.

For guidance on how to go about building ontologies, there is no better place to look than database design. Of course, most of this is now a solved problem in databases, but exposure to the database community benefits expert systems workers, as much by showing them the correct attitude towards design, and by

showing them how many of the things that are problematic in expert systems are commonplace in database. Most important of all, they can absorb the professional attitude and recognition of the importance of detail, and of making detail explicit and complete that is part and parcel of the database world.

A second lesson, which concerns a more current area of database research, can be drawn from interest in *heterogeneity* and *interoperability*. At present there is a good deal of interest in so-called *Knowledge Sharing*: for example there is a very large project in the U.S.A. known as the Knowledge Sharing Effort (Neches et al, 1991). This knowledge sharing work can benefit very greatly from understanding - and adapting and exploiting - database work on interoperability and heterogeneity. I shall return to this later.

4. What Database Can Learn From Expert Systems

The traffic is not, however, all one way, and there are things that have been thought about in the expert systems community that can benefit databases. This is particularly so if we look at some of the areas of current database research.

Since many of the original problems of databases have been solved, attempts have been made to make databases even more useful by adding extra facilities to them: here I shall consider three areas; active databases, databases making use of incomplete information, and distributed databases which may need to reconcile inconsistencies in their information.

Active databases, which are able to unleash a procedure when a data item is accessed, offer significantly increased functionality. They do, however, have some technical problems to solve, in terms of the preventing the active mechanisms from going out of control, and in terms of the semantics of the procedures. This is, however, precisely the mechanism that is at the heart of expert systems, and the solutions that have been advanced in that area need to incorporated into thinking about active databases.

As databases become more ambitious, a need to be able to reason with incomplete information arises. For example, if we have two interoperating databases, one may contain attributes that the other lacks, and we may wish to do better than simply ignore the information that relates to these attributes. Reasoning with incomplete information has been a primary concern of expert systems since their inception, and while it would be too generous to say that the problems have been solved, many solutions have been advanced, and many of the pitfalls have been identified. Expert systems work in this area represents a valuable body of experience that the database community should not ignore.

Finally we should consider inconsistent information. If we have data replicated at a variety of sites, it will inevitably get out of step and require periodic reconciliation. This reconciliation of inconsistency is one of the chief concerns of Artificial Intelligence, Knowledge Representation and Expert Systems, and there is a vast body of theoretical and practical work in the literature. The problem is enormously difficult, but the database community can cut out much of the painful learning process by drawing on this work.

In general, as the database community strives after more sophisticated functionality, it increasingly addresses the concerns that have always been at the heart of expert systems. The history of expert systems is full of mistakes that can be avoided and techniques that can be adopted, if there is an awareness of this.

5. Combining Database and Expert Systems

Databases and Expert Systems both address the same problem - transforming data into useful information which can be used to support the solving of particular problems. So why should not the two be brought together; the expert system exercising its expertise on data drawn not from the interactive input of a user, but from a database? The idea is not new, and in practice some of the more successful expert systems have used exactly this approach. For example the Retirement Pensions Forecast and Advice System, operated by the UK Department of Social Security (Spirgel-Sinclair 1988) is a shining example of an expert system which has proved its worth in terms of money savings and improved service over a period of years. This system derives many of its benefits from the fact that it is able to get its data from an existing database designed for another purpose.

The reason, I believe, that there are not many more applications of this sort is not the lack of opportunity, but the lack of imagination required to seek out this sort of application. Those with databases which could benefit from such an application have dismissed expert systems as unproven, while the expert systems community have been transfixed by the interactive consultative model, and has been rather blind to the opportunities available.

In this way expert systems might become "just another" application exploiting a database, but there may well be good reasons for implementing the application in this way. Indeed this approach might obviate the need for some of the more sophisticated functionality that some wish to graft onto databases. Realising this potential synergy, however, requires mutual understanding and respect between the two communities, and this will only come about if they interact.

7. Federated Information Resources

Finally if we look to the future, we can try to paint a picture of the information systems of the future. Perhaps the most exciting development in computing over the past few years has been the internet and the possibilities for communications that it opens up. The response in the database community has been to explore the possibility of interoperability, and of designing systems of co-operating, although heterogeneous, databases. In the expert systems community similar developments have been termed knowledge sharing, and the potential for distributed, co-operating expert systems has been recognised. Both of these developments are to be welcomed, but they do not go far enough. What we have is essentially the maintenance of an artificial separation between two types of systems which cry out to be brought together. In other words neither co-operating expert systems, not interoperating databases are enough: both need to be brought together in a unified framework, providing a federation of information resources, embracing databases, expert systems, and other types of system, such as information retrieval systems, as well.

This vision is currently being pursued, and architectures have been proposed for such a federation. Perhaps the best known is the mediator architecture of Wiederhold (1992), but there are other alternatives under active consideration. Another example is provided by the KRAFT project (Grey et al 1997), in which I am a participant.

As yet, however, this work is being pursued by specialist research groups, and the vision it reflects is not pervasive in either the database and expert systems communities. As a result, research in these communities is still pursued as though the two classes of system were distinct, and there is little cognisance taken of their shared future. I believe that this is a mistake, both because it may lead to developments which are retrograde in that they do not lead towards the future integration of these systems, and because it allows opportunities to steer the systems closer together to pass by.

If the future of database and expert systems is to be one of integration and cohabitation, it is important that workers in the two fields are aware of developments in the other field, and of the concerns and interests of their counterparts in the other field. This is what makes a conference like DEXA valuable: it brings the two communities together and promotes their awareness of one another. We should not be blinkered, but open to influences from outside our immediate concerns. This is especially true when we have two communities which are as similar in aim as databases and expert systems, and whose future is likely to be a joint one.

References

Duffin, P.H. (ed). (1988). *Knowledge Based Systems: Applications in Administrative Government*, Ellis Horwood, Chichester, 1988.

Grey, P.M.D. et al (1997). *KRAFT; Knowledge Fusion from Distributed Databases and Knowledge Bases*, In this volume.

Gruber, T.R., (1995).*Towards Principles for the Design of Ontologies Used for Knowledge Sharing*. International Journal of Human-Computer Interaction, Vol 43, pp 907-928.

Neches, R., Fikes, R., Finin, T., Gruber, T., Patil, R., Senatir, T., and Swartout, W.R., (1991). *Enabling Technology for Knowledge Sharing*, AI Magazine, 12(3): 36-56, Fall 1991.

Spirgel-Sinclair, S., (1988). *The DHSS Retirement Pension Forecast and Advice System*, in (P. Duffin, ed) KBS in Government 1988, Blenheim On-Line, Pinner, 89-106.

Wiederhold, G., (1992). *Mediators in the Architecture of Future Information Systems*, IEEE Computer, 25(3):38-49, March, 1992.