Knowledge Based Knowledge Acquisition

Nigel Shadbolt

This work presents a framework for knowledge based knowledge acquisition. The ACKnowledge[1] project addresses the problem of supporting the knowledge engineer in the phases of KBS development involving knowledge acquisition. A lot of research in this area has attempted to provide support for this process. This includes the development of AI programming environments, specific knowledge elicitation tools, machine learning techniques, and knowledge acquisition methodologies.

Despite this research substantial problems remain with the acquisition of knowledge from human domain experts, cases, performance records, and documents. In part, this is because the available tools and methods have not been brought together within a single integrated and principled system.

The goal of ACKnowledge is to provide just such a system or knowledge engineering workbench (KEW). Within this project we have been working on the problem of providing a coherent framework for the integration of knowledge acquisition (KA) methods and techniques (Wielinga and Shadbolt1990, Shadbolt and Wielinga, 1990). We have taken the view that KEW will act as a KBS for KA. This view has a number of consequences for the development of the next generation of knowledge acquisition tools.

One of the most important is the realisation that such a KA KBS will have to be an active, dynamic system. It will have to be a system animated by knowledge about how to conduct KA, how to use particular tools, how to transform betweeen them, how to integrate their results into a consistent and evolving application KBS. The issues will be explored in my presentation, and the results of our implementation of an initial prototype of KEW will be discussed.

As stated, a number of approaches to automated support of the KA process have been developed. These software tools fall into three categories: (i) those which have been developed for specific domains (domain dependent); (ii) those which are computer implementations of particular KA techniques (domain dependent); (iii) and those which are integrated systems for acquisition support. Examples of this last class of tool tend to integrate small numbers of KA methods and techniques. It is the aim of ACKnowledge to integrate a much wider variety of techniques into a coherent system.

In this work we view acquisition as a modeling activity. The use of models is a means of coping with the complexity of the development process. A model reflects, through abstraction of detail, selected characteristics of the object, device or process in the real world that it stands for. A number of model driven approaches abound. These can be distinguished in virtue of what the models represent. Some use models of the overall problem solving architecture to be used in the KBS, others use knowledge about invariant aspects of tasks in the generic application domain. A key research issue in our own work is how these various model driven approaches can be integrated: how does one select the most appropriate model driven methods for any application.

AI Group, Dept. of Psychology, University of Nottingham, Nottingham, NG7 2RD.

We also regard the acquisition process as having a cyclic structure; planning sequences of KA operations, technique selection, technique application, assimilation of the results, and evaluation of the revised KB. Supporting these phases of activity requires distinct types of assistance from KEW. We distinguish between; clerical activities such as editing and browsing, KA tool execution where we run an actual technique, operations on the results of KA execution including knowledge transformation and integration both between and within tools, the provision of context dependent advice and guidance.

As we move through these different levels of support KEW must become more active and directive. The corollary is that it requires more and more substantive knowledge to support these kinds of activity. Let us consider the types of knowledge that could provide a foundation for an active KEW.

The first has to do with knowledge about the KA process. This amounts to the encapsulation of expertise about conducting KA. Advice and guidance about what to do and when. This needs to be context sensitive with respect to the activity cycle postulated above. Such advice and guidance takes many forms but includes; how to plan a sequence of KA activities, what are the cost benefits of such a plan (Shadbolt and Burton 1989), what domain/application characteristics constrain particular kinds of KA activity.

The second source of knowledge has to do with models in the broadest sense of the term. It consists of generic knowledge extracted and formulated into models of problem solving methods, epistemological categories, domain structures and content. Thus the adoption of a model of problem solving for a paricular application can be used to direct the use of tools, and the structuring of the KB. Directing acquisition via templates that specify the kinds of structural and static knowledge that may be present in an application or domain is another example of the use of model based knowledge. This sort of expectation driven acquisition can also arise using the general characteristics of knowledge. For example, whether data is measurable, certain, or polymorphic may be knowledge that can direct the use of a particular acquisition tool.

Knowledge about the KE tools themselves constitutes a third influence able to direct KEW activity. For example, repertory grids (Shaw and Gaines, 1987) make assumptions about data, how it is analysed and the status of results. The knowledge engineer must be made aware of, or in some manner, directed through these issues. Background processes are being formulated that examine the data used and produced by the tool. Such processes offer suggestions or critiques about the nature of results and their possible interpretations. Returning to our repertory grid example KEW may seek to determine whether the dimensions along which elements are to be ranked are ordinal, metric, euclidean etc. Are the elements themselves homogeneous. Such knowledge can be used in very direct ways. A common error in repertory grids is to fail to appreciate that an analysis can be rendered meaningless if one dimension in the grid is treated as a linear scale and another as logarithmic.

Finally knowledge gathered through continuous reasoning by KEW on the results of KA activity are also important. We refer to this as evaluative knowledge. KEW contains knowledge about the analysis of the knowledge accrued. Examples include KB analyses for consistency, coverage, generalisation and specialisation.

Each of these types of knowledge is important in the vision of KEW which we present. A prototype has been built which embodies an initial attempt to incorporate these ideas into a working system. The talk will present in more detail how these various components of KEW are being implemented.

[1]This project is funded under the auspices of the ESPRIT programme (Project No 2576). The project consortium consists of the following institutions; Cap Sesa Innovation, Marconi Command and Control Systems, GEC-Marconi Research Centre, Telefonica, Computas Expert Systems, Veritas Research, Sintef, the Universities of Amsterdam and Nottingham.

References

Shadbolt and Burton (1989) The Expirical Study of Knowledge Elicitation Techniques. SIGART Newsletter, April 1989, No 108.

Shadbolt, N.R. and Wielinga, B. (1990) Knowledge based knowledge acquisition: the next generation of support tools. Proceedings of the European Knowledge Acquisition Workshop, Amsterdam.

Shaw, M. L. G. and Gaines, B. R. (1987) An Interactive Knowledge Elicitation Technique using Personal Construct Technology. In Kidd, A. L. (Ed.) Knowledge Acquisition for Expert Systems: A Practical Handbook. New York: Plenum Press.

Wielinga and Shadbolt (1990) Conceptualisation of a Knowledge Engineering Workbench. Esprit Project P2576, Deliverable 1.4.