Relations between Knowledge Engineering and Cognitive Science: From Import/Export to a Truly Interdisciplinary Knowledge Acquisition Enterprise

Franz Schmalhofer

German Research Center for Artificial Intelligence
University Bldg 57
Erwin-Schroedinger Str.
W-6750 Kaiserslautern

email: schmalho@informatik.uni-kl.de

1. Introduction

Knowledge Engineering is generally known as the field that is responsible for the analysis and design of expert systems and is thus concerned with representing and implementing the expertise of a chosen application domain in a computer system. *Research on cognition* or *cognitive science*, on the other hand, is performed as a basic science, mostly within the disciplines of artificial intelligence, psychology and linguistics. It investigates the mental states and processes of humans by modelling them with a computer system and combining analytic and empirical viewpoints.

Early on, *knowledge acquisition* was known as the activity of making explicit the human knowledge that is relevant for performing a task, so that it can be represented and become operational in an expert system. Knowledge acquisition and the field of knowledge engineering are consequently closely related to human cognition, which is studied in cognitive science. The specific relationship between knowledge engineering and cognitive science has changed over the years and therefore needs to be reconsidered in future expert system developments.

Although knowledge acquisition activities are at most twenty years old, there is already a respectable history with noticeable successes and some initially disappointing failures to be looked back upon. Actually, more progress was made by the analysis of the failures than with the short term successes.

2. Early Knowledge Acquisition

Early knowledge acquisition was supported by knowledge acquisition systems such as TEIRESIAS (Davis, 1978) which were designed as front-ends for existing expert systems (i.e. MYCIN) and knowledge engineers viewed knowledge acquisition as the process of transferring knowledge from a human expert to a program.

After it was observed that humans can hardly express their knowledge in distinct chunks, so that each chunk can somehow be transformed into a rule (or some other syntactically defined structure), which would then do "the right thing" in combination with an already existing expert system shell (e.g. EMYCIN), knowledge acquisition became recognized as "a bottleneck in the construction of expert systems" (Buchanan et al., 1983, p.129): Not the development of the original expert system (shell), but the acquisition of the domain specific rules for that shell turned out to be the tough part in building a fully functional system.

Since some of an expert's relevant knowledge is tacit or implicit (Schachter, 1987), experts cannot directly verbalize all relevant rules. Knowledge engineers therefore concluded that some special psychological method would be necessary in order to acquire the desired knowledge in the form that was needed for rule-based (or other) systems.

For mining the expert's deeply hidden knowledge, various data collection and data analysis methods were subsequently imported from psychology into knowledge engineering (Hoffman, 1989) and respective computer tools were built (Boose & Gaines, 1989). Some of these tools were quite successful in constructing rules for relatively small application domains.

This early knowledge acquisition period was determined by the knowledge engineers who emphasized the full implementation of small scale knowledge acquisition tools over a common and cognitively founded design rationale for the expert system and its knowledge acquisition components.

Knowledge engineering and cognitive science followed two separate research agendas during this period and those slots of the research agenda which were difficult to fill from inside the field of knowledge engineering were assigned to the field of cognition (e.g. supplying the rules for some rule interpreter). The cooperation of the two disciplines thus consisted of quickly importing selected research items (vague ideas, theoretical frameworks or methods) from the other discipline in a relatively unreflected way. The use of repertory grids (Kelley, 1955) in knowledge acquisition is probably a good example of such a type of import/export relation between knowledge engineering and psychology which is one of the disciplines contributing to cognitive science.

While the problem of transferring human expertise into computer programs was (at least partially) solved, it was discovered that the knowledge acquisition problem had been incorrectly stated, right from the beginning. One piece of evidence for that is: Even after the successful construction of an operational rule-base, the meaning of the individual rules remained a mystery (Clancey, 1983; p. 241). The maintenance of larger expert systems was consequently impossible. Since such systems were found to have several other disturbing deficiencies (e.g. brittleness), the definition of knowledge acquisition needed to be changed.

3. Knowledge Acquisition as a Truly Interdisciplinary Task

One of the necessary changes in the definition of knowledge acquisition is already well established: Knowledge acquisition is now understood as a modelling activity where models of the conventional expertise in an application domain and models of the target expert system are to be developed (Breuker & Wielinga, 1989). Unfortunately, the cognitive science issues which have become important for successful knowledge engineering are hardly discussed in this context. The nature of different types of models and their relationship to one another needs to be determined: How should the models of existing or future artifacts (e.g. expert systems) be related to models of natural systems (e. g. human cognition)? Can they be structurally similar or even identical or do they need to be quite different? Since knowledge engineering deals with such artifacts and cognitive science with the modelling of human cognition, the two fields need to intensively cooperate to successfully address the question of the relation between the models. Newell's (1990) assertion of describing human intelligence as a symbol system is equally important for this discussion as Searle's (1981) views about intrinsic intentionality and human commitment.

Another question, where the expertise of cognitive scientists needs to be respected by knowledge engineers, is: What kind of mental models (Norman, 1983) do humans develop about expert systems? How are the mental models of a domain expert, of a knowledge engineer and of the future users of some target system related to one another? What kind of mental models are users capable of and willing to maintain and how can the mental models about different systems be related to one another? How can expert systems play the role of fancy representations, which allow the communication of knowledge between the domain expert and knowledge engineer on the one side and the users of the system on the other side?

Knowledge engineers must finally learn to appreciate that expert systems have to function in the real world in order to become a success in business. Unlike the microworlds, in which knowledge engineers liked to test their rapid prototypes, the real world refuses to be (correctly) represented once and for all time by some formal specification. The future application requirements can consequently only be partially predicted. This basic fact is often ignored. Expert systems must be developed so that new types of inputs can be processed at the time when the system is applied (Schmalhofer & Thoben, 1992). In other words, expert systems must allow for situated applications (Clancey, 1991) and that means that they must be end-user modifiable (Fischer & Girgensohn, 1990). These challenging demands can only be successfully met, when the engineering science and the cognitive and social sciences cooperate with the mutual respect for one another, which is required to make an interdisciplinary enterprise a success.

References

Boose, J.H. & Gaines, B.R. (1989). Knowledge Acquisition of Knowledge-Based Systems: Notes on the State-of-the-Art. Machine Learning, 4, 377-394.

Breuker, J. & Wielinga, B. (1989). Models of expertise in knowledge acquisition. In Guida, G. & Tasso, C. (Eds.), Topics in expert system design, methodologies and tools (pp. 265 - 295). Amsterdam, Netherlands: North Holland.

Buchanan, B.G., Barstow, D., Bechtal, R., Bennett, J., Clancey, W., Kulikowski, C., Mitchell, T. & Waterman, D.A. (1983). Constructing an Expert System. in: Hayes-Roth, F., Waterman, D. & Lenat, D.B. (eds.) Building Expert Systems. Reading Massachussetts: Addison-Wesley Publishing Company, Inc. pp.127-167.

Clancey, W.J. (1983). The Epistemology of a Rule-Based Expert System - a Framework for Explanation. Artificial Intelligence, 20, pp. 215-251.

Clancey, W.J. (1991). Situated Cognition: Stepping out of Representational Flatland. AI-Communications, 4, 2/3, 109-112.

Davis, R. (1978) Knowledge Acquisition in rule-based systems - knowledge about representations as a basis for system construction and maintenance. In: Waterman, D.A. & Hayes-Roth, F.(eds) Pattern-Directed Inference Systems. Academic Press, New York.

Fischer, G. & Girgensohn, A. (1990). End-User Modifiability in Design Environments, Human Factors in Computing Systems, CHI'90, Conference Proceedings (Seattle, WA) ACM, New York (April 1990), pp. 183-191.

Hoffman,R. A Survey of Methods for Eliciting Knowledge of Experts. In: *SIGART Newsletter 108*,19-27, 1989.

Kelley, G.A. (1955). The Psychology of Personal Constructs. New York: Norton.

Newell, A. (1990). Unified Theories of Cognition. Cambridge, Massachusets: Harvard University Press.

Searle, J.R. (1981). Minds, Brains and Programs. In Haugeland, J. (Ed) Mind Design, Cambridge Massachussets: London, 282-306.

Schachter, D.L. (1987) Implicit memory: history and current status. Journal of Experimental Psychology: Learning, Memory and Cognition, 13, 501-518.

Schmalhofer, F. & Thoben, J. (1992). The model-based construction of a case oriented expert system. AI-Communications, 5, 1,3-18.