Project Proposal: Automated Knowledge Engineering

Principle Investigator : Maanda Raudzingana
Supervisor: Dr. Karen Bradshaw

March 12, 2014

1 Project title

Automated knowledge engineering for an expert system

2 Problem statement

Knowledge engineering is a critical process in expert systems development. It is the process by which expert knowledge is integrated into a knowledge-based system. For this reason, a substantial amount of research has been conducted in attempt to come up with more efficient and effective ways to go about this process. However the process of knowledge engineering has inherent difficulties. These include time constraints on the expert's side hence rendering the expert inaccessible. In addition, experts are unenthusiastic and usually, there exists a lack of communication between the knowledge engineer and the expert. Furthermore, it has proved time consuming and tedious to have a knowledge engineer reconfigure the knowledge base whenever new facts and knowledge became available.

3 Objective of research

The research aims to investigate the use of knowledge base construction tools available, their strengths and weakness as well as current trends and issues in knowledge representation [1]. Special emphasis will be placed on the automation of knowledge base construction or knowledge representation using the production system approach. This raises the question: can the knowledge representation phase in expert systems development be automated with the help of modern technologies? The research aims to find out if it is possible to eliminate entirely the need for a knowledge engineer by automating the knowledge engineering process. This requires the automation of processes involved, i.e. vocabulary construction, classification of domain objects and their attributes, discovery of parameter relationships and inter-dependencies, representation of core facts within the domain and finally the presentation of a rule-based knowledge representation that can be used by an inference engine [2]. Therefore, the primary aim is to present an efficient approach to knowledge representation in the development of expert systems by automating the knowledge engineering process.

4 History and background

Expert systems [4] were among the first truly successful forms of AI software [8]. An expert system, which is an example of knowledge-based systems, is designed and developed to emulate the decision-making ability of a human expert in a specific domain of knowledge. In order to achieve this objective, the system has to closely represent the expert's experience in the field. This requires that the expert's knowledge is acquired and properly engineered into a representation that the system can understand and use in the inference stage. These phases, knowledge acquisition and knowledge representation have proved to be major obstacles in expert systems development [5]. In early expert systems, the common trend was, and still is in some modern cases, to have a knowledge engineer present whose task would be to effectively engineer the expert's knowledge into a representation that can later be used by the expert system in the inference stage. Several researchers in this field have expressed interest in investigating the possibility of automating the knowledge acquisition and representation phases. This, among other reasons, is a response to the realization fairly early on that the arrangement of information greatly affects the effectiveness of an application in artificial intelligence. Continued research in these areas would contribute to the evolution of knowledge-based systems.

A knowledge-based system is essentially composed of two-subsystems, namely the knowledge base and the inference engine. The knowledge base is an organized collection of known facts about the subject, expressed as objects, attributes and conditions. After engineering, the expert's knowledge is stored in the knowledge base. The inference engine, at the heart of the expert system is the mechanism that applies rules to search the knowledge base for solutions. It effectively interprets and evaluates the facts in the knowledge base in order to provide expert advice. In doing so, it applies what is known to what is as yet unknown [7].

It can be argued that the effectiveness of an expert system in providing expert advice is largely dependent on the system's knowledge base. The knowledge base essentially houses the knowledge to be transferred to a user in need of expert advice. For this reason, research on more efficient and effective methods of knowledge representation and control strategies continues [4]. Hogeveen et al. [3] demonstrates that two classes of knowledge can be distinguished. These are declarative and procedural or operational knowledge. Declarative knowledge is the facts in a domain whereas procedural knowledge is knowledge of how to use declarative knowledge. Knowledge representation involves formalizing the acquired body of knowledge in a symbolic form that can be interpreted.

4.1 Knowledge representation schemes

Four common knowledge representation schemes are in use today. Semantic networks [4] are used as a means to represent expert knowledge. A semantic network is a model of associative memory. It is composed of nodes and links that are linked to each other. The nodes effectively represent domain objects or events while the links, called arcs, represent relationships between the objects. Graphically, a semantic network takes the form of a tree structure or graph structure. They provide best use when relationships between the domain objects are complex. An apparent benefit of semantic networks is their ability to allow adding of new nodes and arcs at any point in the development process [4].

First order logic has been used as an alternative method of knowledge presentation. Predicate calculus, first order predicate calculus, propositional calculus and Horn clause logic are some of the systems used to achieve declarative knowledge representation [4].

The use of frames [4] as a method of representing facts and relationships in a knowledge base is common. A frame is an artificial intelligence data structure, a group of attributes that describe a given object. Attributes are stored in slots which in turn contain values and procedures that change the values. The procedures are executed when values in slots change. Function-specific routines are executed within a rigid structure.

A more prominent and perhaps the most widely understood technique of knowledge representation is the use of production or rule-based systems [4][5]. The knowledge engineering process results in knowledge represented in the form of production rules [6]. Productions are used to represent relationships in terms of premise-conclusion pairs. Actions are executed provided a given condition or a group of these is true. This method offers flexibility in terms of extending the knowledge base. Production rules have a simple syntax, which renders them easy to modify and update [5]. To express uncertainties, which are common in expert systems, certainty factors are attached to the premises or conditions of rules.

A hybrid approach [2] is one of the options available to a knowledge base developer during knowledge representation. This assumes the simultaneous use of any two of the methods mentioned above. This approach presents the advantage of having the benefits of any of the two methods achieved from simultaneous use. However this approach in some cases presents the disadvantage of having to deal with complexities brought forth by simultaneous use of any two of the methods mentioned above.

5 Approach

The first step will be to gather more information on the subject of expert systems, specifically the knowledge representation part of it. This will be done with the aim of finding out about the different knowledge representation schemes and how they can be used to achieve best results. Research will be done on the different tools available with which to build expert systems, how they can be used together, their strength and weakness as well as a study on the expert systems that were developed using these tools.

The next step will be to analyze the architecture of the already existing expert system. This will hopefully highlight issues that may aid the development of the knowledge representation component, which will be plugged into the expert system. The existing knowledge representation will be studied to determine how knowledge has been encoded and how the inference engine uses the existing knowledge base.

A choice will be made after extensive research has been conducted on the tools to use for the development of the knowledge representation application. A vast array of development tools, software packages and programming languages exists that can aid the development process. These have different strengths and weaknesses. Thorough comparisons of these will be carried out to aid the decision process with regards to which tools to use.

After having made decisions on knowledge representation methods and development tools, the next step will be to commence with the development of the application. This will initially be in the form of a rudimentary prototype. More functionality will be added after obtaining required results.

A choice between two categories of knowledge acquisition methods is available, knowledge acquisition through inductive learning methods, and knowledge acquisition directly from an expert [5] . The latter will be used. An interface, which the expert will use to input new facts into the knowledge base, will be designed and developed. Different approaches on user interface designs for this purpose will be assessed before selection. The use of an interface to aid the knowledge acquisition process would help in making the process dynamic, should the need arise. The inputted facts will be processed and transformed into knowledge, i.e. a symbolic representation that the inference engine will use during inference. The output of the application should thus be a knowledge base or knowledge representation, produced without any human engineering intervention.

The next and final step towards the primary aim will be to link the knowledge representation tool with the expert system, i.e inference engine. Thereafter, testing will be done to assess the effectiveness of the knowledge representation component. New ideas to extend the tool's functionality in this regard will be considered during the lifetime of the project.

If time permits, secondary project objectives will be carried out. These are 1) finding alternative ways to enter data facts into the knowledge base, i.e expert's domain knowledge, 2) making the acquisition process dynamic by effectively using what is already known, i.e. what is already stored in the knowledge base to ask subsequent questions, and 3) incorporating a second knowledge representation scheme therefore effectively presenting a hybrid approach to knowledge representation. A hybrid approach has been proven to yield better results as a result of an increase in inherent benefits [2].

6 Requirements

A machine running an operating system capable of supporting the application tools that will be used during the development phase is the primary requirement. This will be used when developing the prototype application.

A collection of facts on a specific domain will be used as input to the tool. These will represent the expert's knowledge and will later be used by the tool to effectively encode the knowledge.

To aid the testing phase, a copy of the actual expert system software will be needed. This is not crucial as the effectiveness of the knowledge representation tool can be tested without the expert system. However if need to use the system arises, a working internet connection will be paramount as this is a web based system.

7 Project timeline

Proposed deadline	**Activity**
28 February	Submit project proposal
03 March	Build project website
11 March	Seminar 1
30 March	Begin research on tools and techniques
14 April	Analyze existing system architecture
30 April	Select tools, techniques and languages
30 May	Submit literature review
02 July	Begin prototype development
10 July	Design and develop user interface
29 July	Seminar 2
22 August	Complete prototype development
29 August	Complete user interface development
8 September	Integrate with inference engine
11 September	Perform testing
15 October	Short paper submitted
27 October	Seminar 3
31 October	Project hand-in
7 November	Cease research website development
19 November	Research oral examination

References

1. BRACHMAN, R. J., AND LEVESQUE, H. J. *Knowledge Representation and Reasoning*. Morgan Kaufmann, 2004.

2. CHANDANA, S., MAYORGA, R. V., AND CHAN, C. W. Automated knowledge engineering. *Engineering and Technology 39* (2008), 511–520.

3. HOGEVEEN, H., VARNER, M., BRÉE, D., DILL, D., AND NOORDHUIZEN-STASEEN, E. Knowledge representation methods for dairy decision support systems. *Journal of Dairy Science 77* (1994), 3704–3715.

4. MARTIN, J., AND OXMAN, S. *Building Expert Systems : A tutorial*. Prentice-Hall, 1988.

5. MATSATSINIS, N., DOUMPOS, M., AND ZOPOUNIDIS, C. Knowledge acquisition and representation for expert systems in the field of financial analysis. *Expert Systems with Applications 12*, 2 (1997), 247 – 262.

6. PRAT, N., AKOKA, J., AND COMYN-WATTIAU, I. An MDA approach to knowledge engineering. *Science Direct 39* (2012), 10420–10437.

7. SAWYER, B., AND FOSTER, D. L. *Programming Expert Systems in Pascal*. Wiley Press, 1986.

8. VAN DE GEVEL, A. J. W., AND NOUSSAIR, C. N. *The Nexus Between Artificial Intelligence and Economics*. Springer, 2013.