W(h)ither Expert Systems? — A View from outside

Ian White
Admiralty Research Establishment, Portsdown, Cosham, Hants.

Abstract

The paper questions the expert system paradigm, both in terms of its range of application, and as a significant contribution to the understanding of artificial intelligence. The viewpoint is that of the systems designer who must judge the applicability of these methods in imminent and future systems. The expert system paradigm, (ESP for short), is criticised not because it is ubiquitously wrong, but because its range of application appears to be very limited, and much promise is made of its application in areas where its success is likely to be little more than a matter of luck. The paper considers the success in both academic and commercial settings. It is suggested that the contribution of the ESP to the wider ambitions of AI is modest, and to the practical user is still a considerable and largely unquantifiable risk.

1.0 INTRODUCTION

The raw nerve of artificial intelligence was surely hit rather than touched when Hubert Dreyfus drew his analogy between AI and medieval alchemy. His theme underlying the remark was that there was an ambition, akin to greed, which wanted the rewards, but without having to pay too much regard to the possibility of fundamental limits which might render such aims ludicrous. The analogy stands inasmuch as gold is the materialist’s holy grail, whereas the scientist’s is pure intelligence! Dreyfus has achieved, intentionally or otherwise, the role of the loud unwanted conscience of AI. A more practical perception of what the systems engineer can expect now, and might reasonably expect in the foreseeable future from expert systems is difficult to determine. There are no defined limits to what expert systems might do, nor any guarantees of what they will achieve. This question of fundamental limits is a neglected area in much of AI, however developments from mainstream computer science in complexity theory, and the study of physical limits in computing are examples of potentially fruitful areas in this respect.

In this paper the paradigm and the application of expert systems is examined, not from the perspective of the committed AI researcher, but from the more generalist viewpoint of a systems engineer involved in the process of defining future IT systems. It is the demonstrated and potential capabilities of expert systems that this paper addresses. The author does not, unlike the critique of Dreyfus and Dreyfus, assert that the Expert System Paradigm is ubiquitously wrong, but rather that its range of application is very small, and too much promise is made of its application in areas where its success is likely to be little more than a matter of luck.
2.0 THE EXPERT SYSTEM PARADIGM

The Expert System Paradigm, (ESP for short), is succinctly that human expertise in a wide range of problem solving domains can be explicated in the form of a series of rules, extracted from the expert/s, which can be stated in some formalism, and accordingly executed on a computer, with varying degrees of operator interaction. Consequently expert knowledge can be represented by some symbolic representation of a ‘domain of expertise’, and a useful discourse about that domain achieved by applying some form of inference scheme to that representation. This domain is the setting for extracting rules from a recognised expert, or group of experts. This is an immensely subjective process, because the expert’s inner representation of his expertise is quite unknown to us, and often to the expert himself. The elucidation of such knowledge under the interrogation of a ‘knowledge engineer’ is a curious and unscientific attempt at function matching. Accordingly if an expert system result does not match the expert’s conclusions, the rules are adjusted in an ad hoc manner in an attempt to better the performance. This may perhaps be a plausible process for small systems, (say several hundred rules), but is most dubious for large systems with thousands of rules. The method gives no significant insight into basic system issues such as robustness, sensitivity and evolvability. For these larger systems some tools are available for completeness and consistency checking, and orthodox configuration controls can also ease the task of managing any large volatile system. It remains nonetheless a dubious process with little to describe the macroscopic behaviour of such inferencing structures.

A ‘useful’ ancillary feature of expert systems is that they can ‘explain’ their outputs by reference to their (expert derived) rules, which helps reassure the user. It should however be realised that this is not in any sense a proof or certification of correctness, but merely a mechanism for providing plausible explanation. It is nonetheless a useful check on whether the use of an expert system’s output is appropriate for the user’s application context.

The above definition of Expert system is rather restrictive, and developments in Knowledge Based Systems (KBS’s), which include expert systems, cover some rather more structured attacks at embodying human knowledge in machines by, for example, defining contexts, embodying causal models and using taxonomies of inference types, to better structure the representation of knowledge. A few of these systems exist already, but the comments above about understanding what these systems can achieve still apply. Given that most of the KBS activity is still research, we can legitimately postpone the question for a few years perhaps, although the system designer is still left with the need to make projections about capability, and must understand the limits of application of the offerings in the AI market-place.

The range of problems to which ESP is being, or has been applied, appears to be unlimited, including medical and other complex natural system diagnoses, man-made systems analysis and diagnoses, chemical analysis and military tasks such as sensor data fusion, planning, and command and control. More elaborate AI techniques are also being applied in these areas, but nonetheless rudimentary ESP is still widely applied and promoted. Clearly there are still many people who believe in ESP. I do not claim that this representation cannot exist and work, merely that the means of obtaining it, sustaining it, operating on it, and understanding it, are for the majority of real world problems, not yet within our grasp. Although the more complex aspects of knowledge