Competence and responsibility in intelligent systems

TONY MORGAN

SD-Scicon, Pembroke House, Pembroke Broadway, Camberley, Surrey GU15 3XD, U.K.

Abstract. The capabilities of Artificial Intelligence have increased dramatically over the last decade. We can now contemplate the use of automation in tasks which were previously considered the exclusive domain of human professionals. Such a possibility raises new legal issues. This paper argues the need for finding methods of assessing the competence of such systems in order to assign responsibility for their actions.

Key Words: competence, responsibility, intelligent systems

1. INTRODUCTION

The last decade has seen a dramatic increase in the capabilities of artificial intelligence (AI). Early work in the 1950s concentrated on AI as a means of understanding human mental capabilities. While this is still an important area of research, much recent activity has been stimulated by the desire to use AI techniques, regardless of whether or not they are an accurate representation of the equivalent human processes. This engineering viewpoint has produced a number of successful operational systems, mostly of the 'expert system' variety. Current developments are coupling second-generation expert systems with other developments such as genetic algorithms, deep knowledge, and neural networks. To avoid a tedious technical debate about the precise definition of terms, this paper treats all of the newer technologies together under the banner of 'Intelligent Systems'. Although using a wide variety of techniques and implementation strategies, the common feature shared by all intelligent systems is a deliberate attempt to emulate some aspect of human performance.

The systems which are likely to enter service during the next decade will raise some new legal issues. There are two reasons for this. Firstly, intelligent systems capabilities can now extend into areas which have previously resisted automation. Just as previous waves of computing have had their impact on clerical and manual labour, the newer technologies address 'professional' areas (including the legal profession itself). Tasks which have always been considered essentially human activities are now candidates for automation, at least in part. A particular example, used throughout this paper for illustration, is the domain of air traffic control (ATC). While automated support has long been a feature of ATC, the controller's job itself is beyond the competence of the existing data processing techniques. However, continual increases in traffic demand cannot
be met by simply adding more controllers: a law of diminishing returns sets in because the controllers have to spend more time in liaison with each other. Most authorities are now examining the use of intelligent systems to take over some of the functions currently performed by controllers.

The second issue raised by intelligent systems is the lack of hard experience in their use. The older technology of data processing has been in commercial use since the 1950s, and the consolidated experience of thirty years is now firmly ingrained in existing practice. This stretches beyond software issues alone, to cover aspects such as the way in which contracts are let for the supply of automation facilities. The new element introduced by intelligent systems is capability to deal with information rather than just data. Several of the underlying assumptions in (say) the procurement cycle for new software are nullified by differences in technology. For example, the ‘specify-build-test’ model of software development, which underlies most software contracts, is a poor fit to some types of intelligent system which may be trained, rather than programmed, to fulfil their functions.

2. THE QUESTION OF COMPETENCE

In order to examine the ways in which intelligent systems could fail, we must first understand what could be considered acceptable in such systems. This relates primarily to the idea of competence. We have an unquestioning expectation that human professionals will be competent in their chosen area. The understanding of the need for competence is reflected in our willingness to provide corresponding rewards; either individually (for example, through professional fees) or by acceptance of a special standing for a group providing a less individual service (such as air traffic controllers). In some way, we would like to be sure that any intelligent system possesses a similar competence before entrusting it with our business, or even with our life. Some insight can be gained by examining the characteristics of human competence. We can briefly consider the dimensions along which we might wish to assess the performance of a human professional.

- Accuracy. Producing results which are ‘correct’, judged by some objective standard. This may consist of making an accurate diagnosis in a complex situation, or choosing the right course of action from a large number of possibilities.
- Consistency. Avoiding variations in accuracy against other factors: for example, tiring quickly, or being easily distracted.
- Flexibility. Coping with new or unexpected situations, shifts in the type of tasks performed, or changes in the balance of the workload.
- Presentation. Producing results in a suitable form. For example, in a written report this would correspond to a good structure, clear writing style, attention to grammar, punctuation, and so on.
- Coherence. Performing tasks by an explicable method or path, so that any results are reproducible and can be audited should the need arise.