The Judge and the Computer: How Best 'Decision Support'?

PHILIP LEITH

School of Law, Queen's University of Belfast, Belfast BT7 1NN, Northern Ireland
E-mail: P.Leith@qub.ac.uk

Abstract. One of the important lessons learned from the early expert consultants is that excellent decision-making performance does not guarantee user acceptance. (1981)[1] Version 2.1 is the final release of ASSYST. During the past several years, the Sentencing Commission has informally surveyed probation officers at training sessions to determine the usefulness of the ASSYST guideline application software. On the whole we have found that the ASSYST program is not frequently used. After balancing the program's usefulness with the resources required to keep it updated, we have decided to discontinue maintaining ASSYST. (1996)[2]

1. Introduction

From a period in the early 1980's much of my research effort was spent in arguing against the utility of legal expert systems and legal logic.[3] The effort ceased with the publication of my *Formalism in AI and Computer Science*[4] which argued that the representational formalisms being used by those in AI and computer science were flawed – that they could not, for example, represent the richness of legal knowledge in any useful way. For most of the 1980's my perspective was obviously flowing against a strong tide: logic programming was in the ascendancy and AI still seemed to offer much in terms of practical legal expert systems. However, as I shall make clear in this paper and the quotations above suggest, there has always evidence in the background that there were and are substantial problems with the attempts of AI researchers to persuade users to make use of their systems, even from before the explosion of interest in the application of expert systems. The

---

[1] E.H. Shortliffe, 1981. Shortliffe was the primary designer in the MYCIN medical expert system. See also Shortliffe, 1995; Shortliffe, 1994; Shortliffe, 1992. Shortliffe has been one of those who has been most considerate of user problems in AI. His field has been medicine, but there is little doubt that the discretion given to doctors is relatively similar to that given to judges.

[2] ASSYST web page at *www.ussc.gov/assyst.htm*. This is a rule-based advisory program related to the US Sentencing Commission's task. For a description of this in AI terms, see Simon & Gaes, 1989.

[3] See for example, Leith, 1986a; Leith, 1986b.

[4] Leith, 1990.

thesis of this paper is a restatement of this argument: those who wish to provide judicial decision support systems would be well advised to heed any lessons which might be learned from the problems of user acceptance.

It may be that there are many 'expert systems' which are presently in day-to-day usage. However, given the fact that there is a distinct lack of confidence in any definition of what an 'expert system' is and indeed which programs demonstrate 'artificial intelligence', any discussion of the success of AI and expert systems in general is difficult to conclude. In law though, the situation is clearer: all those 'expert systems' which have been built in the academic environment have *not* moved successfully over to the workplace.[5] Many in AI will not agree with this, but it is a firm position I take which has not yet been proven incorrect by the evidence available – see the Appendix to this article.

In this paper, I want to revisit the arguments against what I described as 'formalist' approaches and suggest that the current debate about 'decision support' is simply the rerunning of old approaches which were found in the expert systems movement of the 1980s. I shall discuss two approaches, the *cognitive* and the *sociological* which can be used in analysis of the judicial process. These are diametrically opposed. One emphasises the techniques which model human processes (i.e. the 'artificial intelligence') whereas the other emphasises trying to understand the social context of the problem. The context of the problem, I suggest, is much more important than modelling human processes. My general argument is that in order to do computing and law properly – whether it be inspired by AI or by traditional data-processing techniques – the most important first step is to understand the problem. If one understands the problem and the problem is solvable, then deciding which technique to use becomes relatively trivial.

2. What is 'Decision Support' – Just a Failure of the Expert Systems Movement?

Informally, it is easy to see what the concept of 'decision support' means. It is an attempt to support decision makers in their day-to-day tasks. However, we are all essentially decision-makers (e.g., deciding whether to take an umbrella to work or not) and there are many such daily decisions which must be made. There is, therefore, something apart from this general support for decision making which is implied. It is:

1. Decision making in a complex environment, and

---

[5] A consistent position taken by AI researchers is to suggest that there are actually working legal expert systems, and therefore that my assertion is incorrect. I deal with this below in the paper's appendix, suggesting that such claims to success should be viewed with more scepticism, and that for those programs which seem to have moved over are usually built with pre-AI techniques and aims (e.g., Juricas from Rotterdam). Typically one hears anecdotal evidence which is not too far from, "I was at a conference and spoke to a friend who had a cousin who once saw a program which was an expert system, and my friend's cousin said . . .". I exaggerate, but only slightly.

2. Decision making which is based in scarce 'expertise'.

These elements are, of course, basic to the idea of the 'expert system' (but do not fully define it) as it was developed in the 1980s and, it can be argued, the concept of decision support system owes its existence to that of 'expert system'. It is not clear from the literature, though, just how or why the idea of expert system was transformed into the idea of 'decision support system'. This transformation is a limiting of the idea of expert system. But what aspects of an 'expert system' are not to be found in the idea of 'decision support system'?

If we look back to the history of AI we can, perhaps, try to find out how the transformation occurred. The lack of success of the *General Problem Solver*[6] model of AI where general solutions to learning and problem solving were to be found gave rise to the interest in complex environments, since these environments appeared to be relatively closed off from the 'noise' of general cognition. The laboratory success of programs which were produced in the late 1970s such as DENDRAL, MYCIN and PROSPECTOR appeared to demonstrate that complex environments might actually be easier to handle than more diffuse domains. That these programs offered expertise which was rare and expensive was seen to be a distinct advantage since it promised that AI would be a useful discipline. The 1980s then became, as we all know, the decade of the expert system.

In the laboratory it is clear that *some* programs which were being designed could be used in a predictive manner. They did, as the terminology of the decade suggested, 'reason like experts'. Certainly MYCIN was the most heavily discussed of the early programs and there is no reason to suggest that – in the very small area in which it worked – it could not help decide which antimicrobial drug should be given. Many other programs were less well analysed. Indeed the *zeitgeist* of the time seemed to be (as can be seen from the papers at AI conferences) that the researcher had to describe what his program was going to do and how intelligent and useful it would be even before it was programmed. Programs were also frequently heavily discussed without, it appears, any link to reality. For example, Duda, a member of the PROSPECTOR team (which was never actually used in practice) wrote:

> The widespread interest in expert systems has led to numerous published articles that cite impressive accomplishments – practical results that justify much of this interest.
>
> Having been primary contributors to the PROSPECTOR project, we are particularly sensitive to comments about that program, such as the one that appeared in a recent book review in this journal that referred to "... PROSPECTOR's hundred million dollar success story". Unfortunately, this particular statement, which is similar to others we have encountered elsewhere, has no factual basis.[7]

---

[6] Newell & Simon, 1972.
[7] Duda, 1985.

DENDRAL's team even suggested that the lack of feedback they got from users was an indication of successful use (rather than – as most producers of programs would realise – a lack of use of the program) by writing:

> Many persons have used DENDRAL programs . . . in an experimental mode. Some chemists have used programs on the SUMEX machine, others have requested help by mail, and a few have imported programs to their own computers . . . . Users do not always tell us about the problems they solve using the DENRAL programs. To some extent this is one sign of a successful application.[8]

And MYCIN, too, when it was moved over onto the hospital wards went unused.

It is important to concentrate clearly upon this history. There have been many hundreds of programs which have been produced as 'expert systems' and which grew from the optimism created by these three programs, MYCIN, DENDRAL and PROSPECTOR. Though early, they were important in defining the mood and direction of much work in the 1980s. But, did these programs make the move from the laboratory to the world of user? My own belief is that these programs were not successful in their move from the labs: and, others, following through the lists of working expert systems, have found little truth in the claims[9] of 'real-life' use.

Though some in AI and law continue to argue for the success of the expert systems concept, it is clear that many researchers have moved on and accept that there has been a failure. Case-based reasoning is an example of the move away from expert systems in the 1980s mode (i.e., heavily rule-based) towards one which is much, much closer to the Jurimetrics[10] model of analysis of cases and prediction according to factors in the case. With case based reasoning, exemplars are analysed by hand and coded according to the existence of factors deemed as potentially important – these are then used to compare the case in hand. Just why such techniques are described as being based in AI is difficult to ascertain, unless one sees them as arriving through the history of expert systems.

Though the term expert system has been abandoned by many, the *cognitive* flavour of that type of research remains: the aim to replicate the reasoning processes and thinking of the expert. Though this cognitive flavour remains, many researchers now prefer to talk about 'decision support' being their aim, rather than the creation of an artificial expert.

What has actually happened here? Is there a conceptual move? It is certainly not clear from the literature, but I suspect that what has happened is that the early realisation by, for example, Shortliffe that the difficulty in the research work *lies in getting users to want to use systems* has been slowly gaining ground. The expert system, providing its 'immortal' knowledge to non-expert users (as was claimed in the first batch of 'popular' literature) has never been a go-er (particularly in the

---

[8] Buchanan, 316, 1981.
[9] See for example, Östberg, 1988.
[10] See, Baade, 1963.

legal area). The second best goal, then, has been to provide neo-expert systems which are directed at *experts*. And, in order to put distance between the earlier failure and the future, hoped for success, the terminology has been changed to 'Decision Support System'. But will change in goal and terminology be sufficient to ensure user acceptance?

3. Problems with the Cognitive Model

The cognitive model which has driven AI has been one which is intrinsically psychological and individualistic. From the earliest days of AI when it was involved in – for example – GPS, the attempt has been to locate the area of interest in the thinking processes of the individual. When expert systems arrived, this cognitive model continued: rule representation was seen to be the means of using and explicating knowledge – for example, when an expert system was said to have an 'explanatory' mechanism, what was really meant was that it could list the rules which had been triggered. Explanation was seen to be wrapped up in the model of rules, which is a psychological model.

There are obviously advantages in this psychological model. It is relatively clear what the object being targeted is – professional problem solving, say. It is also closely linked with a highly developed experimental discipline (psychology and its subset, cognitive psychology) so that – if need be – there is much expertise in testing and designing empirical research. And, given that AI has been closely linked to the computational metaphor during its history, the benefits of having dialogue with the wider psychological movement should help overcome a potential blinkering effect of the computer – that is, seeing everything from the perspective of how models can be computerised.

In the event, the advantages which might have accrued from AI's psychological model have not been particularly evident. The targeting has been, perhaps, too focused with an extremely unbalanced view of, as in this case, professional knowledge. AI has also not been particularly adept at borrowing from the empirical expertise of psychology: the only experiment has been to produce a program and suck it and see; a 'sucking and seeing' which has sometimes been highly superficial. And, the computational model has not been ameliorated in any way by the contact with the broader field of psychology. In many ways, there has been antipathy between mainstream psychology and those in AI.

But even if AI had paid more attention to wider psychological approaches and techniques, would it have made much difference? I think not. It seems to me that the problem which would not be answered by the psychological approach would be that of defining the context of a computer problem, so that a solution can be achieved. This is primarily one of *systems analysis*, rather than psychology. The systems analyst has to look at the wider context of user need, function and role before any proper decision can be made about how best to design a program. The psychological approach doesn't do this. Rather it takes as a given the individual

and tries to understand/model what that individual does. When applied to AI, this would mean that after analysis and repackaging of 'knowledge' or whatever one wants to call it, the expectation is that individuals will want to use a representation of the original behaviour pattern. Why should they? Has there been any stated need for this? And, may there not be another way to better achieve the same aim without simply remodelling one behavioural mechanism[11]?

We have much evidence that the professional has much control over how he fulfils his tasks. We can see from the literature on implementing judicial or court-based systems that there has been a particular context which has allowed this to be supported. Most frequently, it appears that the reasoning has been that the judges have seen political advantage in helping develop a 'decision support system', yet when the system has been produced, have made little use of it. There is certainly little evidence that the judiciary are actively looking for decision support systems which will help them to carry out their functions, and given this lack of desire we should be careful before we begin to assume that judges will accept these systems.

4. The Problem of Judicial Discretion

Since discontinuing writing about AI and law at the end of the 1980's, most of my research has been into the practice of law (and also the use of computers in that practice[12]). I have looked at lawyers in the UK[13] as well as international practice of patent law.[14] The closer I have looked at lawyers in practice, the less useful do most of the AI programs appear. They do not seem to direct themselves to the particular problems of being a lawyer or being a judge and, at worst, appear irrelevant to legal practice.[15] There is, of course, now a huge literature on lawyering and judging which should give insight to researchers into just what it is that they are trying to achieve, but one rarely finds this literature referenced in the research into AI and law. If, however, my assertion that the real problem in designing 'decision support systems' is one of understanding user needs, it seems to me impossible to produce usable programs without understanding the actual legal process as users experience

---

[11] And there is little discussion in the AI literature of which technique is most appropriate – traditional or 'knowledge based' – with many problems seemingly shoe-horned into an inappropriate AI technique. This is not a problem if the aim is development of techniques; however, it may be one if the aim is user acceptance. The conclusion from work on the rule-based Nervous Shock Advisor (Deedman & Smith, 1991), for example, was that "While these weaknesses *[which arose from using rules]* could be palliated to some extent by good rule-base design, they could not be overcome entirely in a system that relied on rules to do everything".

[12] Leith & Hoey, 1998.

[13] Morison. & Leith, 1992.

[14] Leith, 1998.

[15] Deedman and Smiths' Nervous Shock Adviser for example, was developed by a legal prac-titioner but seemingly not a practitioner in this area. Rather, it was chosen because "it seemed manageable". See Deedman & Smith, 1991. The article appeared in a text called *Operational Expert Systems in Canada*. Note that while, despite the title of the collection, the program *was* operational, it was only so in the laboratory.

it. Why do AI researchers believe that success can be achieved without this research link to practice?

In effect, I am arguing that a sociological approach should be undertaken by those involved in AI and law. If they understand the context of the legal process, then the techniques they advocate can be used successfully. If not, their programs will never make the transition from the laboratory to the courtroom.

One of the major problems in designing decision support systems is that the courtroom is a place with much discretion. As recent political debate in the UK and US over minimum sentences etc. shows, discretion is valued highly by judges. Any computerisation project must acknowledge that this discretion exists, and that it is intrinsic to the task of legal adjudicating. Failure to acknowledge it will lead to computer systems being viewed as too formalistic by users and, if labelled in this way, they will certainly not be used. The sociological approach takes this element of discretion as fundamental because it is also a fundamental assertion of a sociological approach to law that the most important element in the legal process is the *actor*.[16] To positivists and others, the fundamental unit is the legal rule or the case or the case and the rule.[17] In the sociological way of doing things, then, rules and cases are not ignored, but are simply seen as artefacts which are produced and handled by actors. The actors are the important element in the situation, because they are capable of interpreting rules and cases in diverse ways, according to the context in which they are to be found. This implies that if we want to handle law by computer, it is a mistake to concentrate upon the rule (or case) since the fundamental source of the rule (or case) is its social use – i.e., whether or how the actor decides to use it or not.

It also implies that since social usage is not necessarily formalistic, it is not possible to say that rules (or cases) can always be represented in any formalism. They may be, and then again they may not be. Some legal rules (or cases) are applied in a highly formalistic manner. We can think of car parking offences which – no matter how unjust the driver may feel in a given circumstance – the system is set up simply to process and fine whoever is ticketed. Attending court and trying to argue that the fine should not be paid will simply be seen as time-wasting. Rarely is discretion allowed. It is almost as though the car owner is expected to take his punishment 'like a man' and pay up – legal liability here is seen as extremely strict. It would be easy to produce a program which modelled this non-discretionary, strict liability system. Unfortunately, for most legal actions, the court system is not so lacking in discretion. These are the ones which are interesting and which should be subject to discussion. These are the ones where facts and rules and cases are sometimes agreed and sometimes disputed, and where the basis for decision is necessarily non-formalistic. Prediction in these situations may be statistical, but cannot be made according to a formal knowledge representation.

---

[16] In the meaning of 'one who performs an action'.

[17] There has been a debate in AI and law on whether rule-based systems alone, or case-based systems alone are the way forward, or whether there should be a combination of the two approaches.

How much evidence exists for suggesting that this problem of discretion is part and parcel of the system – indeed is at the heart of the system? I argue that there is much. In the following sections I shall look at the UK court system. I use the UK system for convenience, but there is a strong argument that most of the aspects of legal process in the UK can be found in other jurisdictions, too. And where the details may match exactly between jurisdictions, the general complexity of situation is, I have no doubt, common to all jurisdictions.

5. The Courts and the Judiciary

There is a substantial amount of investigative material on the court process. From the early studies by ethnomethodologists where a tape recorder was placed in the jury room, through statistical analysis of decisions, to ethnographic studies of courts, we have a huge bulk of material which argues a general point. This point is that, in practice, the legal system differs markedly from that proposed by the traditional 'black letter' law view – one which suggests that legal process is fundamentally rule-based[18] and operates according to strict precedent. There are a number of points of divergence, which to a large degree depend on the position in the hierarchy in the system: the lower down the court scale one goes, the more is there a concentration upon speedy and efficient processing of matters; and the higher up the scale, the more concentration upon esoteric questions of law. The 'black letter' view has – as many commentators have suggested – concentrated upon these upper echelons (the Court of Appeal and the House of Lords in the UK) and neglected what has been happening in the lower courts. Yet despite – or because of – this lack of academic focus, it has been the lowest courts who have managed to make use of information technology – the first courts in the UK to use computers were Magistrates' Courts, for instance.

It is not only that the traditional view of court practice itself has been undermined by these studies, but also we have been provided with a way of looking at what happens to cases before they arrive at court. Thus we have seen studies of the various kinds of 'plea-bargaining' which occur: where defence and prosecution agree lower charges, say, in return for a guilty plea etc. And in the UK, over 85% of criminal cases are processed with a guilty plea. That is, only 15% of criminal cases will involve judicial decision making which is not to do with sentencing. The anecdotal evidence we have from civil cases indicates that of every 100 cases initiated, 10 will go to the courtroom doors and only 1 will be fully considered by the court. This indicates, perhaps, that while the judicial role is important (since it overshadows all the preliminary work on cases and decisions as to whether to proceed) in terms of the importance of judicial decision support to the legal system as a whole, it is not over-riding.

These studies have all led to an appreciation of the amorphous nature of the legal process – that much of the relevant decision making happens out of sight

---

[18] And that these rules of operation are closely akin to formal rules, rather than behavioural rules.

of the judge. Indeed, there is much concern that this process acts against wider concepts of justice: victims and relatives of victims lose their retributive reward when the defendant can successfully plea-bargain (frequently outwith the view of the judge) for a much lesser charge than victims believe to be just.

However, for consideration of judicial decision support systems, we have two clear areas which most AI researchers believe are of interest: that of sentencing and that of consideration of legal rules. The questions must be, just how formalisable are these two areas in practice; and are they viewed as needing support by their potential users?

In the next sections, I want to look at studies into the UK legal system, which have particularly concerned themselves with judging. The arguments from each of these suggest that the nature of judging does not accord particularly well with the traditional view of the judge as a rule- or case-prescribed 'machine'. This will have consequences for my argument on just how possible are decision support systems in justice administration. This argument is that, in order to insert decision support systems into the adjudication process, there is a requirement for two basic factors:

1. There must be a clearly felt need for decision support;
2. There must be a desire by potential users for the information provided by any decision support system;

And further, since all computerisation involves a formalising of tasks into an ordered and coherent 'system', that:

3. The design of a decision support system should not detract from the discretionary elements which are at the heart of notions of 'justice'.

These requirements have not always been met by many of the AI and law programs produced.[20] In many ways, this has been because there has been an element of pioneering or experimentalism about these projects. The aim has perhaps been more to see how far AI techniques could be pushed in the direction of law, rather than how successfully they could be implemented in practice. This is fine, so long as the experimental nature is realised and clearly stated. However, if our aim is to produce working and usable systems, such a methodology is not sufficient. If working systems are desired, then – as the evidence from software failure demonstrates – we have to think carefully about the needs of users.

My assumption is that computerisation involves formalisation. It is clear that much of the research work in the 1980s accepted that formalisation was a requisite for successful operation of systems: Ronald Stamper's LEGOL[21] tried to provide a logical language for law; logic programming attempted to translate legislation into horn clause formalism; and, frame-based formalisms were seen as essential by

---

[20] Though some, e.g., Hutton, 1995 have begun to consider the problem in a much more coherent manner than it has been approached in the past. Whether they see themselves as promoting AI is a moot point. My reading is that they are closer to the sociological approach I advocate.

[21] See, for example, Stamper, 1980.

others. These attempts were not particularly successful, and some have suggested that it is possible to computerise without the degree of formalisation which was considered essential in this previous work. For example, Prakken,[22], Sartor[23] and Gordon[24] have been proponents of a view which suggests that logic can be used in a less positivistic manner.. Prakken has criticised the critics of earlier logical research who suggest that it is necessarily too positivistic. His argument is that though we can use logic to represent knowledge in a logical language, it is not necessary to be bound totally by this logical representation. That is, that other processing techniques (analogical, inductive or abductive reasoning) can be applied to this representation, so that features such as understanding rules and cases within a context can be gained which are not available under the logical approaches used in the 1980s. These techniques, for example as in the work of Prakken and Sartor,[25] can be used to provide tools to encourage reasoning about law by helping construct arguments and also putting into doubt the arguments of an opponent. In effect, if this approach is successful, that logic is being used in a highly social and non-formalistic manner.

I have doubts about this approach. Logicians have, for long, tried to escape the deductive power and/or constraints of first order logic ('fuzzy logic' was an attempt to do this). In logical circles, these attempts have met with much criticism relating to their technical basis (for example, attacks that the approaches are 'not really logic at all' – see Haack for more detail[26]). However, on a more intuitive level we can see that such approaches *may* be useful: the simple act of thinking about something usually helps us to clarify and resolve problems.

More importantly, though, for my argument here is the prosaic fact that we have no real evidence that such non-formalistic approaches actually work in practice – no system has been built operating upon these principles which is being used by lawyers. It may be that such systems are technically possible and needed by lawyers, but we have no evidence of this. In fact, it may be suggested that lawyers do not really need a system to help them construct argumentation models at all: they seem to be perfectly well able to do so without such tools. Further, I may be wrong in this, but my hunch is that such systems would still be too formalistic for practical use.

My essential assumption therefore remains: computerisation imposes a formal structure upon procedures and knowledge which detracts from the social elements (particularly discretion) which are essential requirements of certain judicial tasks. Other tasks can be formalised without problem (court record keeping, for example) or opposition from potential users, but these are not the judicial decision support tasks being discussed in this paper.

---

[22] Prakken, 1997.
[23] Sartor, 1997.
[24] Gordon, 1995.
[25] Prakken and Sartor, 1996.
[26] Haack, 1978.

5.1. THE MAGISTRATES' COURT

The magistrate's court in the UK is the lowest level of criminal court. It copes with a multitude of cases which deal with motor vehicles, petty crime, drunken behaviour, assault, etc. In an important manner, the Magistrates' Court is both a court for dispensing justice, but also for imposing social control. Frequently the judges in these courts are not legally qualified (but they will have a clerk who is legally qualified[27] and who will always be on hand to provide advice to the judges). The Magistrates' Court, in terms of numbers, deals with most cases. It has been the court which has been able to use Information Technology most successfully, in part because there is a processing requirement much like that of the electricity supplier.[28]

In practice it has been the case that sentencing has been diverse and problematical. Evidence for the differences in sentencing patterns in different courts have caused public outcries at various times. However, there has been limited interest in achieving a harmonised system, and there has certainly been no clear call from the magistrates themselves asking for help in making sentencing decisions. According to one Magistrates' Clerk:

> It is difficult to see how fundamental change in an area's sentencing practice can be affected unless a substantial number of the magistrates meet and agree upon it. That prospect is fairly remote. ... The individualised approach to sentencing is very firmly entrenched indeed. Even if "uniformity of approach" were achievable in the magistrates' court, we would still need to ensure that sentencers attached the same degree of importance to all the important variable factors or sentencing disparities would continue.[29]

Turner further points to a survey by the Magistrates Association which suggests that guidelines in sentencing *are* used by magistrates, albeit with substantial alterations in their application – 129 of the 195 surveyed indicated that they made such changes. However, the study was not particularly academic and rigorous in design and, as Turner suggests, "it seems that guidelines are widely used, though how they are applied in practice is another subject requiring investigation". Turner suggests that there is actually a different form of adjudication being applied in the courts: traffic offences are being treated strictly according to guidelines, but the other instances which come before the magistrates are dealt with in a much more individualised manner.

It seems clear, therefore, that though there is a concern about the evenness of sentencing, there appears to be no desire to have a computer system which helps magistrates to achieve sentencing consistency in all areas. Indeed, Bainbridge's

---

[27] Though the legal qualification can be minimal – see Turner, 1992, 198.

[28] See Adams, 1972 for an early view of this. There are now a variety of sources of information on this way of viewing courts.

[29] Turner, 1992: 199/200.

CASE system[30] which was designed to achieve just this kind of support for magistrates went unused. Magistrates are certainly prepared to be guided in a broad manner in most cases, but in some cases they do feel obliged to follow rule-based behaviour – that is, traffic offences. However, it is clear that there is now a move away from using magistrates to adjudicate in many traffic cases – we have seen in the UK the introduction of fixed penalties so that the magistrate is becoming less involved in this kind of decision. Perhaps this raises the interesting question of whether, when a criminal activity is the subject of highly formalised sentencing and strict liability, it is actually necessary to use magistrates justice at all? In effect, when the situation become ripe for computerisation, it is no longer a properly judicial activity.

But for those who have looked at sentencing from the psychological perspective, the complexity is apparent but there are patterns which do show that discretion is attempted in a meaningful way. For example, Lawrence concludes:

> In summary, these analyses reveal the complexity of the network of influences on sentences for a common offence like shoplifting that magistrates mostly see as less serious than drink driving. There is clear evidence of disparity of treatment at all levels of the analyses of the sentencing patterns, yet the patterns do not appear to involve capricious or arbitrary punitiveness. They seem to be genuine attempts by magistrates to achieve sensible sentences.[31]

For those researchers who are concerned with the area of legal rules and cases, rather than sentencing decision support, the evidence we have of the rule-governed behaviour of magistrates suggests that it, too, is problematical for decision support design. In a classic early study of magistrates, Pat Carlen carried out ethnographic study of two courts. The resultant perspective presented is one where the magistrates have a proclaimed rule-governed system, but in fact the rules are used in a more problematical manner. Of the rule system which the magistrates publicly utilise, Carlen suggests it is presented as a form of 'legal absolutism' (and note that the legalism is applied by those who are frequently without legal training or qualification):

> The substantive dimensions of this legal absolutism portray legal processes as being homogenous, unproblematic, external, inevitable, essential and eternal. Strategies by defendants which threaten to reveal legal processes as being pluralistic, problematic, intentional, contingent, phenomenal and temporal are suppressed by remedial routines comprised of strategically interlocked social rules whose tacit components, though suppressed, are treated as being absolutely judicious.[32]

But in practice, applying legally absolute rules is not the task of the Magistrates' Court. The court has the important element of restating society's goals and aims

---

[30] See, for example, Bainbridge, 1991.
[31] Lawrence, 1994.
[32] Carlen, 1976: 128–129.

– that is, in enforcing social control upon those whose goals and aims are not necessarily those of the 'larger society'. For example, drug users do not always follow the reasoning that cannabis use will lead to heroin and the breakdown of society; not all motoring offenders are convinced that society's needs are met by strict liability in, for example, the matter of defective lighting; and drinkers don't always feel guilt about being drunk in public. Accepting the validity of these views would undermine the very system which the magistrates are attempting to impose upon these 'deviants'. As Carlen suggests:

> In legal settings and in other settings more or less serving as agencies of social control, the legitimated goal of theory is not to reveal an alternative, possible society; questioning the inevitability of the social reaction to deviance threatens to do just that. The only discretion to which magistrates will admit is to that accorded them by the formal rules of law. In cases where magisterial discretion is formally prescribed and invoked, implications that such magisterial discretion is open to situational negotiation or legitimation are indignantly repudiated.[33]

Carlen is arguing that there is a very significant area of discretion open to the magistrates. They, however, in order to make their task of social control easier, deny this discretion to be anywhere near its actual reality. Instead, they suggest that the formal rules govern their behaviour entirely. In reality, the formal rules play little part in the decision-making of the magistrate: the defendant in the court must be objectified and squeezed into the relevant classification, in order to be speedily processed, that the court can get on with the next business. The squeezing is carried out by the various professionals involved in the process – court staff, probationary staff and police prosecutors, as much as by the judge. Those who have been onlookers at the Magistrates' Court will certainly have sympathy with Carlen's analysis.

What should designers of decision support systems learn from these kind of studies of magistrates? Perhaps that systems which provide advice on sentencing and rules might be used in a more complicated environment that is first imagined.

5.2. JUDGES IN THE HIGHER COURTS

From a variety of projects we know that judging is an isolated activity – courts are becoming larger in order to become more efficient and the nature of the task is that the judge rarely sees other judges in operation. One study of criminal judges in Philadelphia suggested:

> Nearly all judges interviewed were surprised by the lack of collegiality among their fellow members of the court. There was almost no socializing outside of a few official functions. Each judge seemed to have two or three close

---

[33] Carlen, 1976: 120.

friends among his colleagues but was rather disdainful towards most of the other members of the court. It was surprising to have every judge interviewed comment so negatively upon the ability of his colleagues. The most typical descriptions involved the terms lazy, slow-witted, inadequate, and mediocre.[34]

Rock, in his study of an English Crown Court, too, commented upon the difficulties judges have in communicating their ways of handling problems. The court studied was relatively small, and allowed communication in a way that larger courts did not. One judge in the study suggested that this meant it was possible to:

> know how other people work and how they are thinking. It helps you and stops you doing what used to be the very bad habits in old one-judge courts of developing idiosyncrasies that you don't have the opportunity of noticing. A judge who comes into lunch and says, 'Guess what I did this morning?' - he's done something rather outrageous and the others say, 'What? You can't have done that!' It's very valuable.[35]

It seems relatively unproblematic, then, that judges are keen to find out just what their fellow judges are getting up to: whether it is with the intention of improving themselves or substantiating their negative attitudes towards the others. Communication, it seems, is a good thing. In one way, the move towards sentencing decision support systems appears to be such a form of communication, and might be seen to be welcome. Unfortunately, at all levels of the courts, we do not see evidence for this being seen in a positive light at all. All the information points to the fact that where communication is seen to be one-way (i.e., 'top down'), then the judiciary are opposed to it and see it as negatively impacting upon their judicial discretion.

The history of the separation of powers in government has always been problematic. The judiciary, under this philosophy, are seen as being merely technical implementors of the desires of the legislature. This, of course, has never been the case[36] and there has always been the potential for a battle of wills between the judiciary and the legislature and executive. Sentencing has been a recent example here (with the Sentencing Commission in the US, for example[37]) but there have been other problems, too. For example, the attempt to improve court productivity through judicial case management (this is a factor in the Woolf Report [38]) has been longer lasting in the US than in the UK and has been the focus of unhappiness from the judiciary. Sheskin and Grau in their analysis of the attempts by the Ohio Supreme Court to enforce court productivity through new rules of practice on the

---

[34] Wice, 1981: 165.

[35] Rock, 1993: 183.

[36] Indeed, one of my arguments against logic being able to formalise the British Nationality Act was that it included an 'ouster clause'. See Leith, 1986.

[37] See Tonry, 1992.

[38] See Woolf, 1996. Currently in the UK there is an attempt to move more decision making in the Magistrates' Courts over to the Clerks (who welcome this) for reasons of efficiency. This is opposed by the Magistrates who argue that it is removing judicial functions over to the administrative sphere. See the report in The Times, 10th February 1997.

judiciary, suggested that the lower court judges had tactics available to undermine these attempts:

> although their professional norms precluded collective actions to oppose or change the rules, they were not without defenses. The rules undermined their professional status, but this very status was a weapon against those who sought to rationalize their work. Respect for judges' professional status prevented the Supreme Court from publicly embarrassing or punishing those who failed to meet the established guidelines. Nor did administrative officials challenge the judges' professional integrity by auditing their statistical reports. . . .
>
> Faced with exploding demands for their services and dwindling budgets to provide them, the courts confront a dilemma to which there is no simply solution. No longer free to expand the quantity of services, they must develop new, less costly resources as well as make more 'efficient' use of old structure and functions. But this transformation is not a simple structural reflex to changed circumstances. Rather it is a process impeded by the interests and power of those upon whose work administrative reforms impinge.[39]

The general conclusions which we should draw from these kinds of researches are that though there are problems in judging which the judges themselves are well aware of, the solution is not simply a case of providing some technocratic solution and expecting it to work. The judiciary have historically been highly sensitive to attempts to undermine their discretion, and it must be assumed that any attempt to impose computer-based decision support solutions upon them will simply not work. Hutton et al. are clear about this, when they try to produce a system which is not prescriptive, but descriptive (although there was also clearly a desire to be prescriptive and 'reform' sentencing without, they believed, affecting judicial discretion). They were aware that a threat might exist:

> there was a concern that some judges might perceive the prototype Sentencing Information System as a threat to their independence. Not only might this lead to resistance in the designing of the prototype, but even if resistance could be overcome, judges might be reluctant to consult a Sentencing Information System.[40]

Other researchers – particularly those from AI – have been less receptive to these problems, which can certainly undermine the viability of a project.

If sentencing has been a particular problem, what about the use of rules? Are these to be found to be less problematic when we move up the court hierarchy? The evidence is that they are not less problematic. Certainly in the more esoteric areas of legal practice where law is being fought and constructed by public agencies and private corporations, there is much battling over legal interpretation. The High Court, the Appeal Court and the House of Lords are surely locations where much

---

[39] Sheskin & Grau, 1981: 248–289.
[40] Hutton et al., 1995.

difficult analysis of precedent is undertaken and the balance of precedents must be considered. We see this in the Law Reports. However, these are not really the agents to whom most AI researchers have directed their activities. Of more concern has been the general practitioner, rather than the esoteric expert - the general practitioner is clearly the potential user of case-based reasoning system or rule-based expert systems. Do these general practitioners actually require such systems? It appears that there is little support for the argument that they do, certainly in the UK. We know how little use is made of legal information systems and how rarely is 'academic-style' legal research carried out by the typical practitioner.[41] In my own research into the professional life of barristers [42] it was clear that most provincial barristers (i.e., the vast majority of barristers) are not particularly involved in this kind of research. One commercial barrister suggested the typical view that law was important, but it was not necessary to carry out legal research:

> lots of *[other barristers]* say you never have to use the law you learned at university but that's not my experience .... I use quite a lot of law, largely the kind of basic law that one learned at university ... you have to make quite a bit of reference to contract, equity and that sort of thing. (94)

And one very successful barrister suggested:

> You don't necessarily have to go into much detail. The sort of opinions I'm delivering, they're adequate; they're a fairly generalised view of matters ... enough to allow the clients to make decisions ... Cases where you have to read all the authorities and go into all the arguments are fairly few and far between. (94)

And one arguing that the basic need in a barrister is to be practical, spoke about an academically inclined colleague:

> ... there is a completely different sort of barrister to the one I hope I am ... the one who enjoys playing with the law .... Funnily enough we've got one in chambers ... he loves talking technical points, and all the judges know it ... he's a bright guy ... he knows ten times as much law, a hundred times as much law, as I will ever know but I don't believe he's as practical in his approach and that I believe causes problems .... Who are you serving? ... At the end of the day, technical problems can usually be overcome, You're only delaying the evil day in most cases. (95)

Of course, this is the barristerial view, but barristers are highly attuned to the needs of judges:

> Well, county court judges, if they see someone with more than a couple of authorities in front of them, they just turn off. High Court judges know the law usually, anyway, and it is a case of teaching your grandmother to suck eggs. It is occasionally a bit insulting to High Court judges. (100)

---

[41] See Campbell, 1976 for a view of provincial solicitors.
[42] Morison & Leith, 1992.

Clearly, the picture being painted here is different from that presented in most AI descriptions of the proposed locations of their AI and law programs. Here the picture does not accord with the more academic view of the importance of research and precedent. Law is important, but in a much more routinised manner and all barristers in this study suggested that their client's case would not be served by too much emphasis upon legal argumentation.

The consequent of this is, of course, that if legal research is not of primary importance to barristers, it can only be because the judges do not see it as important. And if they do not see it in this light, then it is unlikely that they will have a need – as users – for a rule- or case-based decision support system.

The general tenor of findings in *The Barristers' World* has been accepted by barristers, and I believe it expresses the situation which the vast majority of barristers find: they may well wish to actually present legal arguments in court, but the judges are rarely prepared to accept them. The aspect of the legal process with which the barrister becomes most involved (and hence, the judge too is most involved with) is the negotiation over fact. Just why most judges are not predisposed to legal persuasion is not of primary interest here: it is sufficient to conclude that it is not likely that they will feel a need for a decision support system which is grounded in legal rules or cases.

6. Conclusion: What's Wrong with Traditional Decision Support Techniques?

The presumption of AI proponents is that the computer offers a better way of doing things. We can see this from the very early days of expert systems: MYCIN had the worthy aim of attempting to dissuade doctors from prescribing incorrect antibiotics, a practice which has been leading to the lessening of effectiveness of these antibiotics against infection. And in legal AI, too, the general assumption is that the computer will help users to better understand and practice law or, in the case of sentencing, to sentence in a more rational manner.

This is a large presumption and seems to have developed without any proper analysis of exactly what is wrong with the more traditional techniques of 'decision support' – for example, training, education or the use of books. Training and education of sentencers has certainly been considered by many[43] to be an effective way of allowing judges to communicate with each other about their sentencing practices. There seems to be no real criticism of this method, and we can conclude that it is not unhelpful. And books – particularly textbooks – are used by most lawyers to provide a grounding in particular points of law, and are well regarded in practice. Yet, what is it about books which proponents of AI see as such a failing, that they feel a decision support system which might have only 10,000 words is qualitatively better than a textbook which contains 150,000 words? Why is less better?

---

[43] For example, the *Judicial Studies Board* in England and Wales.

It seems to me that this lack of investigation of the failures of traditional decision support methods is just another example of the failure to investigate the social context of the fields in which AI programs are supposed to be effective.

Some 400 years ago we saw a similar situation to the one we have seen over the past two or three decades.[44] Then, it was the introduction of print technology which led a group of advocates of the new technologies to advocate new ways of handling information: in much the same way that today's AI researchers believe that the new technology arising from the computational metaphor will radically alter and improve the handling of information and knowledge, those early pioneers worked with a spirit of hubris attacking the traditional disciplines with relish and restating their knowledge in the form of ramist logic. The attempts – as we now know – were unsuccessful and ramist logic diffused into the intellectual ether and is now considered a deviant period in the history of ideas. I can't help but feel that a similar fate awaits today's AI. Certainly, unless it learns to understand what judges do and what they need, the attempt to aid them will be unsuccessful.

Appendix: What do we Mean by 'Successful Implementation'?

Discussions about the success of AI usually involve citation of hearsay evidence concerning some program that someone heard about; was mentioned as an aside during a conference presentation; has been advertised; or is described in a journal article by the author of the system. Unfortunately, hearsay evidence is not really good enough either in court or in science, and means that many of the claims made for AI are unprovable without further investigation.

It may be that the criticisms in this paper are invalid and that – around the world – there are a large number of legally based decision support systems in use. If so, there should be little problem in enumerating these in detail, describing their usage statistics, length of service, user views, the novelty of technique etc. If there continues to be little detailed consideration, the sceptic is justified in suggesting that 'successful implementation' is more difficult than AI proponents suggest.

There are a number of good reasons why the evidence which is provided in a hearsay context is not useful in a proper consideration of 'successful implementation'. We can suggest some of the reasons under the following rubrics:

1. NOT EVERY ARTICLE PURCHASED IS USED (THE 'GADGET' CRITICISM)

It is well known that much purchased software is never fully used. An office may have AI software packages which were bought – perhaps for evaluation – but were not found to be suitable for their intended task. This is particularly the case where software which lies towards the user's fringe needs is concerned: an advert may have encouraged purchase, or indeed the software may have been available at no cost, and the user has simply acquired a copy, tested it and found it wanting, for one

---

[44] See Chapter 4, Leith, 1990 on Pierre Ramus and his logic.

of many reasons. In this view software is often like the kitchen gadget, purchased but left in the kitchen cupboard.

2. CLAIMS MADE IN ADVERTS ARE NOT ALWAYS TRUE (THE 'SEXY CLAIM' CRITICISM)

There are fashions in software just as in clothes. Marketing people are not slow to link a product with a fashion if they believe that it will boost sales. For example, the notion of air conditioning being operated according to 'fuzzy logic' principles is one which has recently been met: just what air conditioning has to do with the law of the excluded middle is not quite clear, and no doubt the marketing literature glosses over this. The claim, I suspect, is being made without much linkage to reality. There are also many claims of this sort found in the selling of software.

We should be wary of software which is advertised in a certain way (e.g., "intel-ligent") and should – as researchers – attempt to validate the claims made against the actual operation of the software. Thus: is an advertised 'expert system' really what we would call an expert system?

3. PROGRAM WRITERS DON'T ALWAYS RETURN TO CORRECT OPTIMISTIC CLAIMS (THE 'FORGET THE FAILURE' CRITICISM)

The journals of AI and computer science are full of suggestions that programs constructed by the author of the article are being used. There is usually very little discussion of who is using this, or how often, etc. I have tried on several occasions to follow up some of these programs (particularly the ones giving advice in legal areas) but have found it extremely hard to locate users. Usually it is impossible to do so. Recourse to the program author provides a later and more measured analysis describing how the program did not really enter full use. Authors do not hide this information, but one must approach them to discover it: they rarely follow up their original article with an article about the program's failure. This is not surprising: authors are interested in improving their work and want to get onto the next step. However, readers who take the original article as entirely factual are being led to incorrect conclusions.

4. THE DIFFERENCE BETWEEN A TRADITIONAL PROGRAM AND AN AI PROGRAM IS NOT ALWAYS OBVIOUS (THE 'IT WOULD HAVE WORKED ANYWAY' CRITICISM)

It is welcome to see a program which is successful, but even here we must adopt a cautious and scientific approach: programs for which substantial claims are being made require substantial evidence to support these claims. For example, how do we know that the program only operated through the AI techniques being described? Programs are difficult objects to analyse, and there is an obvious need for program

authors to take particular care about describing exactly how their program operates. Some, of course, in legal AI have done this with absolute integrity, but it would be difficult to believe that all AI researchers routinely did.

5. NOT EVERY COMPANY MAKES A PROFIT OUT OF EVERY ITEM (THE 'BECAUSE SOMEONE SOMEWHERE SELLS IT DOESN'T MEAN IT WORKS' CRITICISM)

Companies which deal with AI software do not always need to make a profit from all or some software. It may be that they are simply small part-time businesses run on a shoe-string selling software which is evaluated by users but never fully utilised (see the 'gadget' criticism above). Or they may see the software as fashionable and a means of attracting customers (the 'sexy claim' criticism) to other non-AI software. Because a business exists, does not mean that all its products are successful.

References

Adams, E. (1972) *Courts and Computers*, American Judicature Society, Chicago.

Baade, H.W. (ed.) (1991) *Jurimetrics*, Basic Books, New York.

Bainbridge, D.I. (1991) *CASE: Computer Assisted Sentencing in Magistrates' Courts*, at BILETA Conference, Warwick.

Buchanan, B.G. and Feigenbaum, E.A. (1981) 'Dendral and meta-dendral: their application dimension', in Webber, 1981.

Buchanan, B.G. (1986) 'Expert systems: working systems and the research literature', *Expert Systems* 3.

Campbell, C. (1976) 'Lawyers and their public', in MacCormick, N. (ed.), *Lawyers in their Social Setting*, W. Green, Edinburgh.

Carlen, P . (1976) *Magistrates' Justice*, Martin Robertson, London.

Cramer, J.A. (1981) *Courts and Judges*, Sage, London.

Davis, R. (1984) 'Amplifying expertise with expert systems', in Winston, P.H. and Prendergast, K.A. (eds.), *The AI Business: Commercial Uses of Artificial Intelligence*, MIT Press, Cambridge, Mass.

Deedman, C. and Smith, J.C. (1991) 'The nervous shock advisor: A legal expert system in case-based law', in *Operational Expert Systems in Canada*, Persimmon Press, Elmsford, New York.

Duda, R.O., Hart, P.E., and Reboh, R. (1985) Letter to the Editor, *Artificial Intelligence* 26(3).

Gordon, T.F. (1995) *The Pleadings Game. An Artificial Intelligence Model of Procedural Justice*, Kluwer, Dordrecht.

Hutton, N, Tata, C., and Wilson, J.N. (1995) 'Sentencing and information technology: Incidental reform?, *Intern. J. of Law and Information Technology* 2(3)

Lawrence, J. (1994) *Sentencing Processes and Decisions: Influences and Interpretative Procedures, presented at 2nd Annual Conference*, Institute for the Study of the Legal Profession, Sheffield.

Leith, P. (1986a) 'Fundamental flaws in legal logic programming', in *The Computer Journal* 29(6).

Leith, P. (1986b) 'Legal expert systems: Misunderstanding the legal process', *Computers and Law* (49).

Leith, P. (1990) *Formalism in AI and Computer Science*, Ellis Horwood, Chichester.

Leith, P. and Hoey, M. (1998) *The Computerised Lawyer*, Springer Verlag, London.

Leith, P. (1998) *Harmonisation of Intellectual Property in Europe*, Sweet and Maxwell, London.

Morison, J. and Leith, P. (1992) *The Barrister's World and the Nature of Law*, Open University Press, Buckingham.

Munro, C. and Wasik, M. (1992) *Sentencing, Judicial Discretion and Training*, Sweet & Maxwell, London.

Newell, A. and Simon, H.A. (1972) *Human Problem Solving*, Prentice-Hall, Englewood Cliffs, NJ.

Östberg, O., Whitakeer, R., and Amick III, B. (1988) *The Automated Expert: Technical, Human and Organizational Considerations in Expert Systems Applications*, Teldok, Sweden.

Prakken, H. (1997) *Logical Tools for Modelling Legal Argument*, Kluwer, Dordrecht.

Prakken, H. and Sartor, G. (1996) 'Rules about rules: Assessing conflicting arguments in legal reasoning', *Artificial Intelligence and Law* 4, 331–38.

Rock, P. (1993) *The Social World of an English Crown Court*, Clarendon, Oxford.

Sartor, G. (1997) 'Logic and argumentation in legal reasoning, *Current Legal Theory* 25–63.

Sheskin, A. and Grau, C. (1981) 'Judicial response to technocratic reform', in Cramer, 1981.

Shortliffe, E.H. (1981) 'Consultation systems for physicians: The role of artificial intelligence techniques', in Webber, 1981.

Shortliffe, E.H. (1992) 'The adolescence of AI in Medicine: Will the field come of age in the '90s?', in *Artificial Intelligence in Medicine* 5, 93–106.

Shortliffe, E.H. (1994) 'Health care professional workstations: Where are we now? . . . Where should we be tomorrow?', *International Journal of Bio-Medical Computing* 34(1–4), 45–55.

Shortliffe, E.H. (1995) 'When decision support doesn't support', in *Medical Decision Making* 15(2), 187–188.

Simon, E. and Gaes, G. (1989) 'ASSYST – Computer Support Guideline Sentencing', *in Proc. 2nd Int. Conf. on AI and Law*, ACM Press.

Stamper, R. (1980) 'LEGOL: Modelling rules by computer", in Niblett, B. (ed.), *Computer Science and Law*, Cambridge.

Tonry, M. (1992) 'Judges and sentencing policy – the American experience', in Munro & Wasik, 1992.

Turner, A.J. (1992) 'Sentencing in the magistrates' court', in Munro & Wasik, 1992.

Webber, B.L., and Nilsson, N.J. (eds.) (1981) *Readings in Artificial Intelligence*, Tioga Publishing, Palo Alto.

Wice, P. (1981) 'Judicial socialization: The Philadelphia experience', in Cramer, 1981.

Woolf, Lord (1996) *Final Report: Access to Justice: Final Report to the Lord Chancellor on the Civil Justice System in England and Wales*, copy at *www.law.warwick.ac.uk/woolf/report*.