Artificial Intelligence and Law **7:** 129–151, 1999.

Out of their minds: Legal theory in neural networks

DAN HUNTER
Law School, University of Melbourne, Australia
E-mail: d.hunter@law.unimelb.edu.au

Abstract. This paper examines the use of connectionism (neural networks) in modelling legal reasoning. I discuss how the implementations of neural networks have failed to account for legal theoretical perspectives on adjudication. I criticise the use of neural networks in law, not because connectionism is inherently unsuitable in law, but rather because it has been done so poorly to date. The paper reviews a number of legal theories which provide a grounding for the use of neural networks in law. It then examines some implementations undertaken in law and criticises their legal theoretical naïvete. It then presents a lessons from the implementations which researchers must bear in mind if they wish to build neural networks which are justified by legal theories.

Key words: connectionism, legal philosophy, legal theory, neural networks

1. Introduction

> What we have to learn to do we learn by doing.
>
> Aristotle, *Ethics*

Though a few implementations exist, neural networks have been largely ignored in attempts to model legal reasoning. It is not clear why there is a reluctance to use connectionist approaches in law, though part of the problem may be that the results have been less than entirely compelling. I will argue in the paper that part of the problem may well stem from an inappropriate choice of legal domains and the use of inappropriate legal theories for the type of architectures possible using connectionism.

The aim of this paper is to assess the validity of neural networks from a legal theoretical perspective, in order to show how neural networks might be better built to model legal reasoning. I will question whether currently implemented legal neural networks have been built with an understanding of the legal theories on which they rely. My conclusion will be that they have not, inasmuch as they fail to use the most obviously relevant models of legal reasoning presented by legal theorists. I will also argue that these implementations often use inappropriate data as the basis for the conclusions which they generate. This has in part lead to the lack of success in the use of neural network architectures in legal domains. I want to emphasise however that I am not suggesting that neural networks are necessarily inappropriate in law: I argue only that the implementations to date have been inadequate.

This paper is divided into five sections, including this introduction. The next section will explain the legal theories which are most relevant to neural network development. Then I will examine a number of the published reports on neural network implementation in law. Following this, I will make a series of comments about the legal theoretical pitfalls into which current researchers have fallen, and into which they may continue to fall. In the final section I conclude with a discussion as to how to avoid these pitfalls in future.

I assume a basic familiarity with connectionist theory and methodology, in particular feed-forward neural networks which learn by back-propagation. This is the predominant connectionist architecture implemented in law. Readers unfamiliar with the basics should consult Andersen and Rosenfeld (1989), Fausett (1994), Luger and Stubblefield (1993, pp. 516–26), McClelland and Rumelhart (1988), Winston (1992, pp. 443–69) and Zurada (1992).

Connectionist approaches rely on statistics to generate meaningful results. This leads to an obvious legal theoretical basis for neural networks. For if an inappropriate legal model is used which does not match this assumption, then the network will be flawed from the outset. This is in fact largely what has occurred with the implementations to date in law. To understand why this is so, we need to look at the theory behind the use of statistical methods in law.

2. Legal Theory and Statistical Approaches

> Our whole knowledge of the world hangs on this very slender thread: the regularity of our experiences.
>
> Luigi Pirandello, *The pleasure of honesty*, 1917

In this section I review the work of legal theorists whose research focuses on the use of statistical methods in law. Since neural networks use statistics to derive their conclusions, a legal theoretical examination of statistics will provide a foundation for my assessment of neural networks in law.

Statistical methods can be suitable for law depending on the purpose for which they are used and the model of legal reasoning adopted. The *prima facie* concerns of lawyers with statistical techniques are that:

(1) They provide no normative basis for decision-making; and
(2) They fail to capture any element of the abstract reasoning of judges.

As to the first point, provided that we seek only to define a descriptive, predictive model of adjudication, then there is no concern with this approach. At this point we seek merely to generate a model of adjudication which is consistent with prior decisions and which is an accurate predictor of subsequent decisions. I confine myself to this definition since artificial intelligence and law is inherently a field which deals with descriptive models of legal reasoning and not the normative (deontological) basis for the outcomes of decisions. Though normative/logical reasoning approaches are central to much of artificial intelligence and law, I suggest that generally these researchers seek to use logic-based formalisms to define a descriptive model of adjudication or argumentation. They are not really concerned with 'deontology' as philosophers would use the term.

Therefore, most researchers in artificial intelligence and law should examine how legal theorists view judicial adjudication, and particularly the degree to which judges are constrained by the law in the decisions they make. Naturally this involves a subsidiary analysis about the way in which lawyers predict the outcome of cases, based upon their understanding of the adjudicatory models adopted by a given judge or judges. I will narrow any use of 'legal reasoning' down to but two issues: the degree of discretion or flexibility a judge has in adjudication, and the degree to which the outcome of cases can be predicted prior to the event.

This approach stems from the work of the American legal realists like Gray (1924), Holmes (1881, 1897) and Llewellyn (1931, 1940) who set out to challenge the doctrinal rules on the basis that they were inadequate predictive models of judicial decision-making. The legal realists were equally suspicious of the 'abstract reasoning of judges,' since they recognised that judges very often fail to articulate the actual motivations for decisions; see in particular Frank (1930). *A fortiori* the critical legal theorists – perhaps best represented by the theories of legal reasoning expressed by Kennedy (1986, 1991) and Boyle (1985) – are extremely sceptical of the formalist thesis that judicial decisions are generated by abstract, objective and non-political reasoning.

This broad acceptance of a descriptive analysis of law provides the underlying reason to remain agnostic about the use of statistical methods in law, and consequently to suggest that connectionist approaches may be appropriate to model legal reasoning. Two broad movements have relied heavily on statistical analysis of law: the behaviourists and the socio-legal theorists. The behaviourists began as a movement during the 1960s and 1970s and are distinguished by their use of sophisticated statistical techniques in analysing adjudication. The work of Kort (1963a,b 1965), Lawlor (1963, 1964, 1968, 1972, 1981), Nagel (1960, 1963) and Schubert (1964, 1968) all showed statistically significant correlations in appellate court decisions between the political attitudes of judges and the outcomes of cases in their courts. At times the correlation was strained or the model presented was inadequate to predict a likely change in the basic doctrinal analysis. However their work showed that even in the unusual arena of appellate court decision-making one could use statistical models to predict decision-making. These lessons have been learnt by at least one artificial intelligence and law researcher, Popple (1993, pp. 23–7), though interestingly his work is not related to connectionism.

More current than behaviourist theories is the growing importance of the work of the socio-legal theorists. This can be contrasted with the behaviourists' work in that it seeks not to examine judicial decision-making at appellate court level but rather is concerned with sociological, empirical and interactionist studies of lower court decision-making. Works such as the seminal examination of McBarnett (1981) and those of the Centre for Socio-Legal Studies at Oxford, show that lower jurisdictional courts follow distinct patterns in decision-making which are both different from the appellate court decision patterns and more importantly often differ from the stated doctrinal legal position. These theorists all reject to a large degree the kind of abstract black-letter analysis of law in favour of seeing how law actually operates in practice. This work also relies heavily on statistical analysis.

Both the behaviourists and the socio-legal theorists provide an argument for adopting a statistical model of legal reasoning, at least as a descriptive analysis. It is strange then to note that researchers in neural networks have been largely ignorant of these approaches. Instead they have relied on traditional doctrinal analysis in testing whether their systems match 'the law' (*viz* what textbook writers say the rules are), rather than seeing whether neural networks may be a useful descriptive (behaviourist or socio-legal) model of adjudication in a domain. In the next section I will briefly examine a number of implementations of neural networks in law. These reports are representative of many others, for example Rose (1994), Rose and Belew (1989, 1991), Bench-Capon (1993), Birmingham (1992), Fernhout (1989), Groendijk and Oskamp (1993), Raghupathi et al. (1991) and Terrett (1995). I will note where other implementations differ. The systems examined are the works of Warner (1991, 1992, 1993), Bochereau et al. (1991), Philipps (1989, 1991), Hobson and Slee (1994) and Bench-Capon (1993). After examining these representative systems I shall analyse the limitations of neural networks in light of the discussion of legal theory.

3. Legal Neural Network Implementations

> No man's knowledge here can go beyond his experience.
>
> John Locke, *An essay concerning human understanding,* 1690.

The systems described below share a number of similarities in approach. I shall therefore group them together under two headings: using neural networks to mimic rules and problems with the training set.

3.1. USING NEURAL NETWORKS TO MIMIC RULES

This is a common feature of neural networks in law, but is perhaps best exemplified by Hobson and Slee (1994). They sought to index a set of theft cases by use of a neural network. In this sense it can be perceived as analogous to SCALIR (Rose and Belew 1989, 1991; Rose 1994). Using neural networks for retrieval of legal information presents no real legal theoretical concern, because the system is not replacing any feature of the human reasoning process. It might be argued that information retrieval involves reducing the amount of the choice of the judge or lawyer, and therefore is actually a form of reasoning. However this is *a fortiori* true of books and all forms of manual or electronic information retrieval. It is unduly fastidious to worry about this issue and hence I shall assume that neural networks like SCALIR, which only retrieve legal information, pose no difficulty.

Since the purpose of Hobson and Slee's systems seems to be the same as SCALIR it would seem that the same comments would apply to their work. However there are some salient differences, most notably, Hobson and Slee's use of a feed forward network rather than SCALIR's interactive network. Hobson and Slee used a number of different feed-forward networks, with anything from eleven input neurodes to the final version with thirty-nine. However, in each network, there were two output nodes, representing whether the defendant was found guilty or not guilty. The authors fail to state as much, but the use of a feed forward network with output nodes of 'guilty' or 'not guilty' makes this a network for predicting the outcome of cases, and not merely for retrieval. Thus, it is appropriate to examine Hobson and Slee's implementation and their underlying assumptions about legal reasoning.

The most problematic feature of the implementation involves the training set: there are only six genuine cases in the training set and they are all leading cases decided by appellate level courts. The use of them to create reasoning neural networks causes enormous problems. Leading cases are, by their very nature, the exceptional cases. Cases are taken on appeal, and become leading cases, because they provide difficulties which make them worthy of the expense of the higher court chance. These difficulties – whether they be alternative interpretations of fact, uncertain or conflicting law, or disputes between law and morality – all serve to create an exceptional case. Neural networks generalise using statistical correlations. Thus, using exceptional cases is unlikely to show us anything except that the network can classify on even ridiculous bases.

The use of leading cases as the basis for law is not just confined to neural networks. In fact, it has been the basis for all 'black-letter' doctrinal law teaching since law teaching began. In this sense, Hobson and Slee might well argue that they are only doing as law teachers or lawyers have been doing for years. There are two responses to this rejoinder, one relating to law teaching and the other to law practice. First, the behaviourists, socio-legal theorists, legal realists, critics or postmodernists are steadily gaining ground on canonical teaching. Hence, the days are long gone when one could apply the doctrine without further discussion or justification. Secondly, and perhaps more importantly, while doctrine might be practically beneficial in symbolic legal expert systems, it is next-to-useless in sub-symbolic approaches. Symbolic legal expert systems can reason with doctrine and deduce consequences from the mixture of rules and facts. They can be used by lawyers to see the rules on which the law is supposed to be based. I would argue, in keeping with Critical legal theories, that doctrine cannot be applied without close thought as to what we are doing in representing it in this form. But at least in symbolic systems we can query the system as to the basis of its decision.

However, a sub-symbolic system is unable to explain the basis for the decision and certainly cannot do so by reference to 'rules'. The neural network's 'rule base' is simply the collection of statistical weights which it has learnt. This statistical information is of little use to the user asking why the network advises as it does. Hence, neural network developers would be better advised seeking not a doctrinal analysis of the law but rather a descriptive analysis. Since neural networks can generate statistical correlations very quickly and generate predictions from these correlations, they are excellent at modelling the law as it is, rather than as black-letter doctrine would have it. Neural networks then seem well placed to give both practitioners and legal theoreticians accurate predictions about the likely outcome of the case.

The final concern with Hobson and Slee (1994) lies in the choice of input nodes. In the final version of their network, the authors chose to include input nodes such as:

> Has the accused picked wild mushrooms, flowers, fruit or foliage for reward
> Has the accused picked wild mushrooms, flowers, fruit or foliage not for reward
> Has land been appropriated by a trustee, personal representative, attorney, liquidator or otherwise

And so on.

Much of this is probably drawn from the rule-based doctrine of theft, and I need not repeat observations made above about the dangers of mimicking rules. However, there is another place from whence these inputs might be drawn, thought the paper is unclear on this point. They might well be single reference points for one case: that is, 'Has the accused picked wild mushrooms, flowers, fruit or foliage not for reward' may be drawn from the unusual fact situation of one particular case. If this is so, then the network is merely performing a kind of pattern matching – something at which a neural network is adept, but which hardly marks it out as anything special since simple Boolean retrieval could perform the same function. Thus, in designing the inputs we need to give consideration to the process of matching, and guard against a kind of simple one-to-one comparison.

A similar emphasis on using neural networks to mimic rules is found in Warner (1989, 1991, 1992, 1993). These were amongst the first implementations of neural networks in law. Warner (1989) suggested the creation of a legal expert system based largely on logic, but later papers, Warner (1990, 1992, 1993) focussed on neural networks. In Warner (1990) he discussed a proposed rôle for neural networks in law but sidestepped any of the issues in neural network creation and maintenance. Instead he argued that a neural network could be viewed as an interesting way of mimicking deontic logic. He proposed that law was a parallel process rather than a serial one, and hence neural networks based on a parallel architecture were appropriate to model law. However, no suggestions were made as to the form that such a neural network might take, and no justification was made of the claim that law was somehow ‘parallel’.

Warner (1992) went some way towards implementing a system. Using Brainmaker, a commercial neural network simulator, Warner implemented a four-layer feed-forward network. It operated on the legal sub-domain of consideration in contractual or quasi-contractual matters. It had 100 inputs, two hidden layers of 25 nodes each, and an output layer of eight nodes. The inputs represent a range of facts that related to the presence of consideration in the contract. Warner (1992) describes one fact situation for consideration as relying on the existence of ‘detriment’ and ‘bargain’. That is, the presence of activations on ‘detriment’ and ‘bargain’ neurodes in the input layer should create the necessary activation at the ‘consideration’ node in the output layer.[1]

Warner (1992) has a number of serious legal theoretical difficulties. First, it seems that Warner judges the success of his system on the basis of whether it arrives at the same conclusion that a production rule or logic program would come to in the same circumstances. This can be seen in the example that Warner gives: activations on the input nodes representing ‘detriment’ and ‘bargain’ leads to an activation on the output node for ‘consideration’. Alternatively, in production rules we might code this as:

IF detriment AND bargain THEN consideration

The use of neural networks to replicate logic based formalisms has a distinguished provenance, but it was rejected when neural networks began operating effectively as sub-symbolic reasoners rather than symbolic-replicating reasoners. From a legal theoretical perspective all this example shows is that it is possible to make neural networks restate doctrinal rules if one only provides enough examples of the rule in the training set. Since neural networks rely on statistics, this seems an inappropriate use of the technology – surely the work of the behaviourists or socio-legal theorists would be a better starting point.

Strangely, for someone who relies so much on rules, Warner (1993) examines the nature of law and parallelism. Warner (1993, p. 136) argues that law is a parallel process that is inherently misdescribed by the emphasis on serial conclusions from a fact-law pattern. In unequivocal language he states:

> While our language dictates a sequential description of the [legal reasoning] process, the process *is in fact* parallel. Many aspects of the problem resolution process *are* carried out simultaneously. The problem domain *is* defined by the

[1] As an aside, Warner provides little explanation of the design decisions in creating his system: its description is confined to a short paragraph and one footnote, and did not explain why he chose to have two hidden layers, and any effect that the extra hidden layer might have had. As Mital and Johnson (1992, pp. 262–3) note, anything which can be performed by multiple hidden layers can be done by a single hidden layer. There appears no need for the additional layer from a technical perspective.

> initial statement of the problem. That initial problem *is* then resolved into a number of issues ... the solution to which *will be* sought within the problem domain *utilizing a subsymbolic paradigm that is not rule based*'[2]

However, when it comes to assessing his simplified network he once again reverts to the simple rule-based analysis. He suggests that in order to test whether the network is training properly, we must observe whether a particular input node has a strong influence on the outcome. Warner (1993, p. 140) then states:

> This observation can then be checked against our conventional rules governing (the domain) to legitimate, or give us some degree of confidence in, the Network's (sic) output.

Once again we find rules at the heart of the creation and testing of a neural network, which once again is not justified nor justifiable using current legal theories.

We see rules again used in the final example of this section (Bochereau et al. 1991). The researchers here discuss a system called Neurolex, which was a neural network extension of MAIRILOG, a suite of expert systems intended to aid mayors in their decision making relating to municipal law. No indication is made as to the form of neural network used. It sought to predict the decisions of the Conseil d'Etat, a major French jurisdictional body which investigated by-laws promulgated by mayors. The network comprised 49 input nodes (grouped in four categories), four hidden nodes and two output neurodes representing annulment or confirmation of the by-law. The network was trained on 331 cases, and checked against the known outcome of 47 separate cases.

While the system can be praised for its separation of training set and validation set, and the relatively large number of genuine cases it uses, the developers once again resort to rules to verify the output of the network. Correctness of the system was, like Warner's flawed approach, tested against the doctrinal rules. Yet the domain used appears much more like the small-scale domains of the socio-legal theorists where the doctrinal rules may very well not exist. In this sense then the Neurolex system may actually be a better descriptive model of the domain than even the developers recognise. I will return to this point when reviewing legal theory in section four.

3.2. PROBLEMS WITH THE TRAINING SET

Another problem that we find with current implementations is the reliance on inappropriate training set data. For example in Hobson and Slee (1994) there are a number of concerns with their training set and the input nodes which they use. First,

[2] References omitted, my emphasis. Warner appears very sure of this conclusion. The references omitted do not justify his confidence. They comprise a contrast (cf) to E.H. Levi, *An Introduction to Legal Reasoning* (1949, Chicago: University of Chicago Press), and a reference to a basic text on connectionism, not law.

the training set is very small, comprising only twenty-six cases. Neural networks classify and generalise by statistical analysis, which relies on correlations between variables in a large data set available. To classify using the twenty-six cases in this training set seems statistically troublesome. This is a feature common to many implementations and will be examined in section four.

The second major problem related to the training set issue involves the use of hypothetical or prototypical cases We see this in Warner (1992). Warner mentions that he has 100 input nodes, but makes no mention of the number of cases used in the training set. In order to train such a net effectively, thousands or tens of thousands of cases would be needed. No discussion is given on where these cases stem from. If, as seems likely, many were hypothetical, we strike the problem of the neural network operating as a quasi-symbolic reasoner. The reason for this is simple, and stems from the way in which hypothetical cases are generated. In order to 'test' if the neural network trains correctly on a training set, the typical way of generating hypothetical cases is to 'reverse engineer' the cases from a rule. That is, create a series of supposed inputs ('What would happen if we have detriment and bargain') and then, using a rule (IF detriment AND bargain THEN consideration), specify the outcome (consideration exists). In this way, we train the neural network on the basis of a rule. In law this approach presupposes the existence of such a rule in the first place. Since legal theories differ over the primacy of rules in law, and most recent theories reject doctrinal rules altogether, training a neural network to generalise a doctrinal rule tells us little about the efficacy of neural computing as a model of legal reasoning.

We see this also in the Hobson and Slee (1994) training set. Only six cases of their twenty-six are genuine. Here there is the dual problem of hypothetical cases, and very few cases. This was not the case however with Bench-Capon (1993). The system worked in the domain of a fictional welfare payment for people visiting a spouse in hospital and assessed eligibility of payment. *Prima facie* this appears to be a good domain in which to work, given that it conforms with the approach of the socio-legal theorists discussed in section two above. Bench-Capon tested a number of alternate network architectures, all of which were feed-forward networks, trained using back propagation. He varied the number of hidden layers, using 1, 2 or 3 layers of hidden nodes, and assessed the convergence of each after training.

Each network was trained on 2400 cases. These cases were generated by a small Lisp program: the output from the program represented what would be the result (true or false) of the presence or absence of six necessary and sufficient conditions. For example, these conditions included whether the person applying was of pensionable age, whether the person had paid pension contributions for five years, whether the person was a spouse of patient, and so on.

In connectionist terms this was methodologically the most sound implementation of any neural network we have seen to date. This is so because of the large number of cases used in the training set and the different architectures tested. This is to be contrasted with the work of Warner, and that of Hobson and Slee.

Table I. Philipps' training set

Inputs (existence of mother, son, daughter)	Output (ratio of estate paid to mother: to son: to daughter)
(1, 1, 0)	(1 : 2 : 0)
(1, 0, 1)	(2 : 0 : 1)
(0, 1, 1)	(0 : 1 : 2)
(0, 0, 0)	(0 : 0 : 0)

Further, Bench-Capon was able to show that noise inputs are disregarded and that the network would classify cases. Finally, he showed, somewhat surprisingly, that the neural network actually classified on the basis of five inputs, and regarded one of the necessary conditions as irrelevant. He makes much of this in the paper.

For all that this is methodologically the most appropriate approach it must be said that it is not a paper about *legal* neural networks at all. Instead it is about the use of neural networks to simulate necessary conditions, since the training set and the verification set were simply drawn from rules. Again, these rules were simply the doctrinal basis of the (fictional) legal area. Once again, this tells us little, if anything, about whether neural networks are appropriate to model legal reasoning, though it may be relevant to the discussion whether neural networks can model rules. But as the realists and socio-legal theorists suggest, law is not really about rules at all. Or at least, law is not about the kind of rules that Bench-Capon (1993) uses. This problem is intimately tied up with his use of hypothetical cases.

Finally, the systems described in Philipps (1989, 1991) are interesting, not so much for the implementations but because they were produced by a legal theorist, Lothar Philipps of the Institut für Rechtsphilosophie und Rechtsinformatik at the Ludwig-Maximilians-Universität, München. Philipps implemented two simple neural networks, one relating to a Roman law hypothetical, and one on damages assessment in car accidents.

The Roman law example focussed on a hypothetical testator who was mortally ill and whose wife was pregnant. His will provided:

> (I)f a son is born to me let him be heir in respect of two thirds (of my estate), let my wife be heir in respect of the remaining part; but if a daughter is born to me, let her be heir to the extent of a third; let my wife be heir in respect of the remaining part. Philipps (1991, p. 988)

In the way of these parables, the wife delivers of twins: a boy and a girl. Philipps trained a network with a training set as follows:

When confronted with the hypothetical (1, 1, 1) – that is, mother lives, son born, daughter born—the neural network generated the solution (2 : 3 : 4)—mother receives two shares of estate, son receives three shares, daughter receives four shares. Philipps suggested this was a reasonable solution, particularly given that the result could not have been merely a recognition of a prior case, since (1, 1, 1) was not used in the training set. Philipps argues that the use of a neural network here is interesting because it is a means of providing equilibrium where there is conflict, which, he argues, is the best way of approaching the problem of distributive justice. This approach can perhaps be viewed as a kind of Nozickian view of justice (Nozick 1974), to be contrasted with a Rawlsian view (Rawls 1972). I seek here not to criticise Philipps' notion of 'Equilibrium as Justice', but rather to point out that there are other views of justice that would not sit happily with the use of neural networks to resolve fundamentally insoluble problems. Though Philipps does not mention it, the same issue arises in the hypothetical examples of the wife delivering of identical-sex twins, triplets or a stillborn child. The example that Philipps gives is perhaps less an example of the benefits of neural networks but rather a demonstration of the insolubility of such problems by any mechanism: an interpretation that Philipps (1991, p. 991) does note.

Philipps (1991) then reports on a neural network for damages assessment in traffic accidents. This network was also, it appears, a feed forward network which learnt by back propagation. A small training set of ten 'prototypical' cases was used, and the ten inputs included elements such as 'Both drivers using the same lane', 'Driving on the autobahn' and 'Stopping without apparent reason'.

There are two matters worthy of mention here. First, Philipps uses 'prototypes', that is, cases which he suggests 'define the problem set'. In this way he argues that he is able to train the network with a very small number of cases. The view that we take of this depends very much on the view that one takes of the idea of prototype cases in law. Are they the vital determinants of all other cases that Philipps assumes? If we take a radical postmodern legal theory, the notion that one can define the entire corpus of law by reference to ten cases seems troubling. However, it may be possible to suggest that within a given interpretative system, in this instance the German crash-and-bash litigation system, these prototypes define the range of results which a court or a litigator might generate. However, this is a large assumption and one which should be tested empirically.

Secondly, Philipps notes that he is working in a domain where there might be some conflicting inputs. For example, if two people were driving in the same direction and in the same lane at the time of the accident, then an actual court will almost certainly decide for the driver in front. If the same facts hold true, but we have also an input that the driver in front turned into traffic from a parking space, then the court should rule in favour of the rear driver. The effect on the network of such inconsistency was that the outputs generated for a single fact situation depended on the rate at which the system learnt. The range of outputs swung widely. The range was from +0.97 to +0.61. To suggest that these figures do not matter, is to similar saying that almost certain (0.97) is the same as just over two-third chance (0.67). This sits ill with the notion that there is a small range of reasonable answers to a problem of this kind. Further, it seems anomalous that the only reason given for such a large variance is the rate at which the neural network is supposed to learn. Philipps concludes by arguing that perfection is impossible and so a system based on a statistical simulacrum of justice or equilibrium seems reasonable. This is a strange claim to make for someone who purports to care about justice.

Having examined the specific implementations, what are we to make generally of the legal theoretical issues in neural network development? In the section that follows, I draw some general conclusions about the nature of neural networks in law and the dangers of using inappropriate cases in the training set.

4. Legal Theoretical Criticisms

> (The passage from the book by Jose Luis Borges) quotes "a certain Chinese encyclopedia" in which it is written that "animals are divided into: (a) belonging to the Emperor, (b) embalmed, (c) tame, (d) suckling pigs, (e) sirens, (f) fabulous, (g) stray dogs, (h) included in the present classification, (i) frenzied, (j) innumerable, (k) drawn with a very fine camelhair brush, (l) *et cetera*, (m) having just broken the water pitcher, (n) that from a long way off look like flies". In the wonderment of this taxonomy, the thing we apprehend in one great leap, the thing that, by means of the fable, is demonstrated as the exotic charm of another system of thought, is the limitation of our own, the stark impossibility of thinking *that*.
>
> Michel Foucault, *The order of things*, preface xv.

Let us look at the two basic issues identified in the discussion of implementations above: the nature of neural networks and the nature of the training set. I will argue in the first sub-section below that neural networks perform pattern matching which is not analogical reasoning. Further, they should not be seen as a metaphor for law in their parallel nature. To draw high-level comparisons between neural networks and the whole legal system is pointless. Then in the second section, I will argue that neural networks can only be appropriate in modelling legal reasoning if the training set of cases they use is comprised of a large number of low-level, commonplace cases. There are a number of problems with the training sets of implementations to date, which I will examine, and from which draw a number of conclusions about the implementations.

Let us look first at the nature of neural networks.

4.1. THE NATURE OF NEURAL NETWORKS

Neural networks have performed well in classification and recognition of 'objective' sensory data. The most obviously successful field has been visual perception, where pattern recognition works extremely well, see Winston (1992). The statistical basis of the paradigm means that it is very good at making correlations between a new pattern and a previously trained one. In law we might expect neural networks to perform similar tasks equally successfully. Neural networks can perform useful pattern matching by 'recalling' a previous identical case to the one at bar, see for example Hobson and Slee (1994). Further, a network's statistical basis allows it to recognise associations between related cases by increasing the weighting on the links which correspond to these cases. In this way the neural network can perform classification-type processing. If presented with a sufficient number of cases which contain similar attributes and values, together with relevant outcomes, the neural network can classify these cases as being of one type. For example, in Warner (1991)'s contract law network, 1000 cases containing the inputs of offer and acceptance with outcome of valid_contract would create a strong classification regime. Then, when presented with a new case (offer = yes and acceptance = no), it will 'recognise' this case as falling within this classification and return the given outcome (valid_contract = no).

The question remains however: Is legal reasoning simply a process of pattern recognition or classification? Many theories of legal reasoning suggest that it is not, any more than it is a simple process of symbolic deduction. If legal reasoning were classification, then each case would fall within a particular classification or set of classifications, and a given result would follow. Since any practising lawyer can present alternative arguments on either side we know this is not all there is to legal reasoning. We do not have to choose between (Hart 1961)'s reduction of hard cases to core and penumbra, or the critical legal theorists' view of indeterminacy (Kennedy 1986, 1991), or other more radical postmodern views (Fish 1982, 1983). We simply need to recognise that in law generating only one answer is insufficient, whether that answer is generated by logical deduction or by pattern matching. Implementations like Hobson and Slee (1994) and Warner (1991, 1992, 1993) fail to recognise that classification is not reasoning.

The problem with classification is not the only difficulty resulting from claims that neural networks can perform legal reasoning. Related to the preceding discussion, there seems to be an assumption by some researchers that these systems can perform analogical reasoning. By analogical reasoning we mean the ability to retrieve precedents, adapt them better to fit the current case, and draw conclusions from the relationship, see Ashley (1992). Can neural networks perform analogical reasoning? Hobson and Slee (1994) have, as an assumption of their discussion, that their neural network system is to be used for case-based reasoning. Philipps (1991, p. 998) shows a diagram that indicates a number of 'conclusions drawn by analogy', when these conclusions appear to be simply based upon classification.

To say that a classification or generalisation process produces results based on analogy is misleading. It falsely suggests that the neural network can modify its conclusion by a process of case adaptation, when in fact no implementation has yet attempted this. For Hobson and Slee (1994) or Philipps (1991) to claim that their simple neural network performs analogical, or even simple precedential, reasoning so is misleading. Other researchers who use neural networks simply for classification do not present such problems.

If these were the only concerns about the use of neural networks then we could be sanguine about their future development in legal domains. However other broader claims are made. Recall that Warner (1993, p. 136) suggested that the nature of neural networks is such that it is not only useful for legal reasoning, but, more grandly, that it somehow represents the entirety of law. His is a grand theory of parallelism indeed, singularly lacking in any authority, examples or other backing for the argument presented. Other authors are somewhat more circumspect, though they still advance interesting claims for neural networks. For example, Rose and Belew (1989, p. 140; 1991, pp. 2–3) in their SCALIR project suggest that the legal system contains elements of both localised and global concepts. That is, legal doctrine stems from the interaction of the corpus of judge-made and statute law. As they put it, 'Thus global concepts emerge from the interaction of a large number of localised decisions' (Rose and Belew 1991, p. 2). They then suggest that this bears a striking similarity with the symbolic and connectionist approaches of artificial intelligence research. Hence, they argue, an integrated connectionist and symbolic architecture is appropriate for their legal model and retrieval system.

Rose and Belew (1991, p. 3) argue further that the ability of SCALIR to adjust its weights by back propagation accords with legal realists' view that '... law can(not) be adequately explained by some set of rules or concepts'. I may not take issue with their artificial intelligence formalism, but am concerned about their adoption of such a simplistic view of legal realism. Rose and Belew (1989, 1991) are in fact conflating the myriad strands of all descriptive jurisprudence with one type of statistical analysis. It may be true that statistics has been, and will continue to be, used by some legal realists, many socio-legal theorists, and by the behaviourists, but that is not to say that all realist jurisprudence is limited to statistics. They are in effect viewing all of legal reasoning as a statistical process, rather than as a process of individual cases, interlocking obligations, relationships of power, race, and so on. It seems, at least to me, too simple a model to be accepted without further study. It is a view of legal realism with which many legal realists would be unhappy.

Rose and Belew (1989, 1991) might be accurate in arguing that the use of neural networks in *retrieval* accords with a legal realistic view of law. This seems reasonable, given that their SCALIR system is designed only for document and information retrieval. However, this obscures the vital point that neural networks may not accord with a legal realist view when they are used for *reasoning*. Neural network conclusions are based on a statistical analysis of a training set. Assuming the training set is valid – a question addressed in the next section – the output generated is simply a conclusion based upon a statistical weighting of the importance of the inputs. Some would argue that they cannot suggest that all law is necessarily based upon statistics. Appeals to statistical data have often been rejected by courts, and represent at best only one way of analysing a given legal domain. The point is not to reject statistics altogether, but to recognise that it is but one model of the legal system, and moreover a model that must be used with caution.

There is one final point worth mentioning before turning to the nature of the training sets. We need to consider the inability of neural networks to explain their conclusions. Being non-symbolic, neural networks cannot explain to a lawyer why they concluded as they did. All of their ‘intelligence’ is in the weightings on the links between neurodes. This information is not of a type amenable to symbolic manipulation and hence explanation. The lawyer asking the network why it came to the conclusion it did is going to be very disappointed when it can provide no justification. Therefore, we need to consider how to provide an explanatory function to neural networks, perhaps by integrating them with rule based systems (Stranieri et al. 1994; Stranieri and Zeleznikow 1992), or by deriving symbolic information from them (Gallant 1993).

Warner (1993), Bench-Capon (1993) and Bochereau et al. (1991), suggest analysing the weights of the neural network links in order to derive symbolic information. That is, determine which links are most important in the neural network, and from this information draw tentative rules from the network. Quite apart from the difficulty of performing such an analysis, particularly on a large network, one must ask why they are using an inherently sub-symbolic approach to generate symbolic information. If this is all we seek, then we would be better served relying on statistics *simpliciter* or on logical deduction.

Nonetheless, whatever the concerns raised by the nature of neural networks, they are small compared with the problems associated with the training sets used to train these neural networks. It is to this issue that we now turn.

4.2. THE NATURE OF THE TRAINING SET

The training set comprises the basis for the knowledge of any neural network, unlike symbolic systems where the encoded rules provide the intelligence. Thus, if we are to generate anything of value from a neural network, we must be careful in choosing the information encoded in the training set. Unfortunately, this has been perhaps the greatest failing on the part of the legal neural network implementors.

There are two basic features of the training set which we need to consider – the types of cases in the training set, and the attributes chosen to represent the cases as neural network inputs. Let us assess the nature of the cases which the systems to date have used for training.

4.3. TYPES OF CASES

The first point to bear in mind is the sheer number of cases which neural networks require in order to train properly. This is due to their statistical foundation—it is impossible to adjudge any feature as statistically significant unless it is seen in a vast range of cases. Thus, a neural network needs thousands, or at least hundreds, of cases to learn properly. We have yet to see this in the legal implementations. With a few exceptions, Bench-Capon (1993) and Bochereau et al. (1991), the systems implemented rely on few cases. For example, Philipps (1991) relies on ten cases, while Hobson and Slee (1994) rely on twenty-six. How can one draw any statistically significant inferences from this small training set? Philipps (1991, p. 996) suggests that there are 'prototypical' cases which define the subject matter. He argues that if one uses only, or mostly, these prototypical cases then one can train a neural network to generate the correct answer. This is nothing more than constraining the training set in order to have the neural network mimic a symbolic reasoner.

To illustrate the problems with this type of approach, let us say that we seek to create a neural network to assess whether a driver will lose her licence to drive because she was drunk while driving. It has two inputs (drive and drunk) and one output (licence_loss). Let us say we can train the network using only two prototypical cases:

Case 1: The driver was drunk and driving and lost her licence.
Case 2: The driver was not drunk and did not lose her licence.

If we train the network with sufficient repetitions, it will generate the expected answer. However, if we are relying only on prototypical cases – that is applying the existing doctrinal rules – why use a neural network at all? We could use a production rule system (IF drunk AND driving THEN licence_loss) which is more computationally efficient, and has the benefit of an explicitly symbolic representation, with its associated explanatory features. Tiny training sets which assume a correct doctrinal answer seem to be a misuse of the neural network, and represent the unthinking acceptance of doctrine as the only correct basis for reasoning in law.

There is another aspect which keeps appearing in legal implementations—the use of hypothetical cases. Due largely, it seems, to the need for large training sets and the difficulty of obtaining such large sets in law, implementors have chosen to supplement their training sets with hypothetical cases. Some implementations add hypothetical cases to genuine ones (Hobson and Slee 1994), while others rely only entirely on hypotheticals (Warner 1991); Oskamp et al. 1989 and Bench-Capon 1993). The distinction is immaterial.

'Padding' the training set with hypotheticals seems at first benign, until we consider that these cases are derived from a rule. That is, a rule is specified (for example , IF drunk AND driving THEN licence_loss) and then cases are generated in huge profusion which satisfy the rule. The effect of this upon training the network is, once again, to make the network simulate a doctrinal symbolic rule-based system. That is, the network should be able to induce the rule from the cases. Once again, we see doctrinal rule positivism creeping into our use of neural networks. As I discussed above, doctrinal formalism is a much discredited model of legal reasoning. Since neural networks have the potential to operate extremely well using legal theories based on statistics, it is quite remarkable that researchers persist with naïve doctrinal analysis in justifying their conclusions.

The blind adherence to doctrine in neural networks is not limited to the use of hypothetical cases in the training set. When one looks at the genuine cases used in a system such as (Hobson and Slee 1994) one sees that they use the leading cases of the domain. Since leading cases are, by their nature, exceptional, using them as the basis for statistical analysis is virtually guaranteed to generate poor conclusions. We should try to avoid implementing in a neural network that disease which Frank memorably diagnosed as 'appellate-court-itus'? (Frank 1930) That is, ignoring lower court decisions which are amenable to statistical modelling, and instead concentrating on upper court cases which are not amenable to this type of analysis. Using neural networks intelligently means that we must choose a domain where the descriptive power of the paradigm can be used. These domains will be where there are large corpora of similar cases, and they are likely to be found in the lowest level courts of first instance. These domains, like car accidents, marital dissolutions and work related injuries, are much more likely to give us the basis for meaningful networks than upper court areas like theft, murder, and so on.

Finally, while examining the choice of cases in the training set, consideration must be given to conflict. In particular, the issue arises: how does a neural network deal with conflicting cases in the training set? Traditionally, in building symbolic systems, the way to deal with conflicting cases was to choose one case as better representing the doctrine, and discard the cases which conflicted with it. One can do this with neural networks, but like its symbolic cousins, this approach is almost completely unsatisfactory. What use is a system which cannot tolerate conflict, since this is such a vital and distinguishing feature of law?

Philipps (1991, pp. 992–993) suggests that neural networks can handle conflict in the training set. He says of his neural network that:

> Remarkably, the network also tolerates contradicting learning patterns. The equilibrium attained by the units will be a compromise between the patterns.

Philipps identifies the means by which neural networks handle contradiction: contradictory cases simply lower the weightings on some of the links. So for example, if we have a number of cases which indicate a positive outcome, and one case which indicates the contrary, we will see the weight reduced on the link associating the facts with the positive outcome. The problem arises, of course, where we have one case which is the most important or most recent, and which best represents the law as it is progressing, and a series of old, dated cases which indicate what the law once was. The important case will be overwhelmed by the sheer number of less important ones.

Philipps argues, however, that conflict and landmark cases pose no concern, since the most salient aspect is whether equilibrium is achieved. He argues that striving for equilibrium in conflicting cases is both appropriate and something neural networks do well. His rationale for arguing that equilibrium is a basis for justice relies on the symbol of law being the scales. This is hardly a persuasive argument. As we saw previously his view of 'justice as equilibrium' does not accord well with some of the major theories of justice advanced in the last twenty-five years; theories as diametrically opposed as Rawls (1972) and Nozick (1974) do not rely on notions as morally neutral as equilibrium. However, it is interesting that even Philipps, who argues for equilibrium, notes that neural networks find it difficult to create equilibrium when given contradictory data. He notes that the inconsistency creates a range of possible results for the networks. The results he generates are not derived from any feature of the data, but rather depend almost completely on the rate specified for the learning algorithm. Such an artefact of the implementation seems neither to satisfy the community's expectation of 'justice' nor Philipps' more modest requirement of 'equilibrium'. More work must certainly be undertaken to examine how we can resolve concerns resulting from contradictory data.

4.4. TYPES OF INPUTS/FACTORS

We need also be wary of what input neurodes represent. (Hobson and Slee 1994) chose to include input nodes such as:

> Has the accused picked wild mushrooms, flowers, fruit or foliage for reward
> Has land been appropriated by a trustee, personal representative, attorney, liquidator or otherwise

I have argued that this type of approach is probably relying on the pattern matching ability of the neural network, rather than any type of classification or reasoning. We must therefore not be too impressed with any network which can retrieve cases based on single case pattern matching, where the input nodes represent in effect one case.

Intelligent identification of relevant input nodes can however be a real strength. They may provide weight to alternative legal theories about a given domain. For example in death penalty cases, rather than expressing what the judges *say* are the important criteria in assessing whether the death penalty is appropriate (for example, 'violence' or 'previous convictions' and so on), the system may give credence to what we think might be better explanatory features (for example, what is the race of the defendant and victim). We may find that the neural network generates accurate predictions of outcomes, without reference to the doctrinal basis. This is, of course, a sketch of the approach of the socio-legal theorists, and neural network developers would be advised to see how these theorists approach the analysis of their domain.

We must still be careful. For example, we can use factors to allow us to justify any conclusion. Let us say that we are Kennedy (1986)'s judge, trying a murder case who has decided for political reasons how she wants the decision to come out. Let us also say there are a whole slew of murder cases against this outcome. To overcome the weight of these precedents she might argue that none of them have considered a new factor, say that the murder was committed with salad forks.

Training a network on the old cases will ignore the new factor, and the new case can be resolved by the new factor alone. Thus, any case can be justified by adding a new factor.

5. Conclusion

> The mind is a strange machine which can combine the materials offered to it in the most astonishing ways.
>
> Bertrand Russell, *The conquest of happiness* 1930.

I have argued that neural networks have major problems from the perspective of legal theory. First, there is the problem that statistical correlation and classification is an inadequate model of legal reasoning. It may be appropriate in some domains, but the domain has to be chosen carefully. Unthinking use of statistics will give rise to inappropriate models of the domain. This is clearly expressed by Michelman 1983, p. 201):

> If an elegant theory appears to explain, in the correlational sense, a respectable amount of the variance in a set of observations, there will be some tendency to picture the variance remaining unexplained by that theory as unsystematic, random, impenetrable muck lacking significance – what statisticians call "the residuals". Thus one may be led to think that the distinctive behavioral theorems of economic [or scientific] analysis ... are the only ones capable of making comprehensible the legal phenomena that interest us.
>
> Critics of law and economics (myself included) believe the opposite is true ... (that) which is the "systematic" component and which the "random" is very much in the eye of the beholder .... The critical point is to avoid mistaking an organizing construct for a structural reality that, by defining the possible, limits vision and deadens will.

Further I argued that the very grand claims of neural network researchers that law is a parallel process or somehow analogous to neural networks appears unsustainable. Finally, on the use of neural networks generally, I suggested that connectionism's lack of explanatory capability provides a major barrier to its use in law, though not perhaps a legal theoretical barrier.

I then presented a number of examples where neural networks were built using inappropriate data, such as hypothetical cases, small training sets and poor domain theories. Though these are not legal theoretical bars to neural networks generally, they do represent major flaws with the implementation of neural networks in law. Until these basic problems are remedied we cannot be confident that neural networks can be used in legal domains. It may be that new approaches in neural network design or in fuzzy logic can answer all of these criticisms. But we are yet to see these approaches implemented in law, and we must remain undecided on the merits of these answers.

This paper has examined neural networks in light of legal theory. There are equally cogent issues raised by the above implementations from a purely computer science perspective, see for example (Thagard 1991) However, this is beyond the scope of this paper, but there is scope for significant work to be done on it.

Neural networks may be appropriate for legal domains, but we must give careful consideration to artificial intelligence theory and legal theoretical models before advocating their use. In particular, it is worrying that we are seeing the tacit acceptance of rule positivism again. This seems to ignore the true benefits that a statistically based paradigm such as neural networks can bring to our study of law.

Acknowledgments

Thanks to those who read and commented on previous incarnations of this work: Richard Ingleby, Richard Johnstone and David Hamer (University of Melbourne, Australia); Andrew Stranieri, Mark Gawler and John Zeleznikow (Latrobe University, Australia); Michael Aikenhead (Durham University, UK); Trevor Bench-Capon (University of Liverpool, UK); Lilian Edwards (University of Edinburgh, UK); Patricia Hassett (Syracuse University, USA); Danièle Bourcier (Sorbonne-Pantheon, France); and to attendees at the Seventh Dutch Conference on Legal Knowledge Based Systems, Universiteit van Amsterdam, Netherlands, and at the First French-American Conference on Artificial Intelligence and Law, Syracuse University, USA. Thanks also to the anonymous referees of this journal for their suggestions.

Finally, thanks to Dr Herchel Smith, whose endowment of my fellowship at Emmanuel College made this work possible.

References

Andersen, J.A. and Rosenfeld, E. (eds.) 1989. *Neurocomputing: Foundations of Research*. Cambridge: MIT Press.

Ashley, K.D. 1992. Case-based reasoning and its implications for legal expert systems, *Artificial Intelligence and Law* **1**(2), 113–208.

Bench-Capon, T.J.M. 1993. Neural networks and open texture. In *Proceedings of the Fourth International Conference on Artificial Intelligence and Law*. Amsterdam: ACM Press, pp. 292–297.

Birmingham, R. 1992. A study after Cardozo: *De Cicco v. Schweizer*, noncooperative games, and neural computing, *University of Miami Law Review* **47**, 121–145.

Bochereau, L., Bourcier, D., and Bourgine, P. 1991. Extracting legal knowledge by means of a multilayer neural network application to municipal jurisprudence. In *Proceedings of the Third International Conference on Artificial Intelligence and Law*. Oxford: ACM Press, pp. 288–296.

Boyle, J.D.A. 1985. The anatomy of a torts class, *American University Law Review* **34**, 1003.

Fausett, L.V. 1994. *Fundamentals of Neural Networks: Architectures, Algorithms, and Applications.* Englewood Cliffs: Prentice-Hall.

Fernhout, F. 1989. Using a parallel distributed processing model as part of a legal expert system. In A.A. Martino (ed.), *Pre-Proceedings of Third International Conference on Logica, Informatica, Diritto*, Volume 1, Firenze: IDG, pp. 255–268.

Fish, S. 1982. Working on the chain gang: Interpretation in law and literature, *Texas Law Review* **60**, 551–567.

Fish, S. 1983. Wrong again, *Texas Law Review* **62**, 299–316.

Frank, J. 1930. *Law and the Modern Mind*. Brentano. Republished 1963, Garden City: Anchor Books, 1970, Gloucester: Peter Smith.

Gallant, S. 1993. *Neural Network Learning and Expert Systems*.

Gray, J.C. 1924. *The Nature and Sources of Law*. 2nd ed, New York: Macmillan.

Groendijk, C. and Oskamp, A. 1993. Case recognition and strategy classification. In *Proceedings of the Fourth International Conference on Artificial Intelligence and Law*, Amsterdam: ACM Press, pp. 125–132.

Hobson, J.B and Slee, D. 1993 Rules, cases and networks in a legal domain, *Law, Computers and Artificial Intelligence* **2**(2),119–134.

Hobson, J.B. and Slee, D. 1994 Indexing the Theft Act 1968 for case based reasoning and artificial neural networks. In *Proceedings of the Fourth National Conference on Law, Computers and Artificial Intelligence*, unnumbered additions, Exeter: Exeter University Centre for Legal Interdisciplinary Development.

Holmes, O.W. 1881. *The Common Law*. Boston: Little Brown & Co; reprinted London: Macmillan, 1968.

Holmes, O.W. 1897. The path of the law, *Harvard Law Review* **10**(8), 457.

Hunter, D. 1994. Looking for law in all the wrong places: Legal theory and legal neural networks. In H. Prakken, A.J. Muntjewerff, A. Soeteman, and R. Winkels (eds.), *Legal Knowledge Based Systems: The Relation with Legal Theory*. Lelystad: Koninklijke Vermande, pp. 55–64.

Kennedy, D. 1986. Freedom and constraint in adjudication: A critical phenomenology, *Journal of Legal Education* **36**, 518.

Kennedy, D. 1991. A semiotics of legal argument. *Syracuse Law Review* **42**, 75.

Kort, F. 1963a. Content analysis of judicial opinions and rules of law. In G. Schubert (ed.), *Judicial Decision-Making, International Yearbook of Political Behavior Research*, Vol. 4. New York: Free Press of Glencoe, pp. 133–197.

Kort, F. 1963b. Simultaneous equations and boolean algebra in the analysis of judicial decisions, *Law and Contemporary Problems* **28**(1), 143.

Kort, F. 1965. Quantitative analysis of fact patterns in cases and their impact on judicial decision, *Harvard Law Review* **79**, 1595.

Lawlor, R.C. 1963. What computers can do: Analysis and prediction of judicial decisions, *American Bar Association Journal* **49**(4), 337–344.

Lawlor, R.C. 1964. *Stare decisis* and electronic computers. In G. Schubert (ed.), *Judicial Behavior: A Reader in Theory and Research*. Chicago: Rand McNally, pp. 503–505.

Lawlor, R.C. 1968. Personal *stare decisis*, *Southern California Law Review* **41**, 73–118.

Lawlor, R.C. 1972. Fact content of cases and precedent – A modern theory of precedent. *Jurimetrics Journal* **12**(4), 245–270.

Lawlor, R.C. 1981. Analysis and prediction of judicial decisions, in R.P. Bigelow (ed), *Computers and the Law: An Introductory Handbook* (third edition). Chicago: American Bar Association, pp. 81–84.

Llewellyn, K.N. 1931. Some realism about realism – Responding to Dean Pound, *Harvard Law Review* **44**(8), 1222.

Llewellyn, K.N. 1940. The normative, the legal and the law-jobs: The problem of juristic method, *Yale Law Journal* **49**, 1355

Luger, G.F. and Stubblefield, W.A. 1993. *Artificial intelligence: Structures and Strategies for Complex Problem Solving*. 2nd ed, Redwood City: Benjamin Cummings.

McBarnett, D. 1981. *Conviction: Law, the State, and the Construction of Justice*. Oxford: Clarendon Press.

McClelland, J., Rumelhart, D., and the PDP Research Group. 1988. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Cambridge: MIT Press.

Michelman, F.I. 1983. Reflections on professional education, legal scholarship, and the law-and-economics movement. *Journal of Legal Education* **33**, 197–209.

Nagel, S.S. 1960. Using simple calculations to predict judicial decisions, *American Behavioral Scientist* **4**(4), 24–28.

Nagel, S.S. 1963, Off-the-bench judicial attitudes. In G. Schubert (ed.), *Judicial Decision-Making*, New York: Free Press of Glencoe, pp. 29–53.

Nozick, R. 1974. *Anarchy, the State and Utopia.* Cambridge: Harvard University Press.

Oskamp, A., Walker, R.F., Schrickx, J.A., and van den Berg, P.H. 1989. PROLEXS divide and rule: A legal application. In *Proceedings of the Second International Conference on Artificial Intelligence and Law*, Vancouver: ACM Press, pp. 54–62.

Philipps, L. 1989. Are legal decisions based on the application of rules or prototype recognition? Legal science on the way to neural networks. In *Pre-Proceedings of the Third International Conference on Logica, Informatica, Diritto*. Florence: IDG, p. 673.

Philipps, L. 1991. Distribution of damages in car accidents through the use of neural networks, *Cardozo Law Review* **13**, 987–1000.

Popple, J. 1993. SHYSTER: *A Pragmatic Legal Expert System*. PhD. diss., Department of Computer Science, Australian National University.

Raghupathi, W. et al. 1991. Exploring connectionist approaches to legal decision making, *Behavioural Science* **36**, 133–139.

Rawls, J. 1972. *A Theory of Justice*. Cambridge: Harvard University Press.

Rose, D.E. and Belew, R.K. 1989. Legal information retrieval – A hybrid approach. In *Proceedings of the Second International Conference on Artificial Intelligence and Law*. Vancouver: ACM Press, pp. 138–146.

Rose, D.E. and Belew, R.K. 1991. A connectionist and symbolic hybrid for improving legal research, *International Journal of Man-Machine Studies* **35**(1), 1–33.

Rose, D.E. 1994 *A symbolic and Connectionist Legal Information Retrieval System*. Hillsdale: Lawrence Erlbaum.

Schubert, G. (ed.) 1964. *Judicial Behavior: A Reader in Theory and Research*. Chicago: Rand McNally & Co.

Schubert, G. 1968. The importance of computer technology to political science research in judicial behavior, *Jurimetrics Journal* **8**(3), 56–63.

Stranieri, A., Gawler, M., and Zeleznikow, J. 1994. Toulmin structures as a higher level abstraction for hybrid reasoning. In *Proceedings of the Seventh Australian Artificial Intelligence Congress*. Armidale: World Scientific, pp. 203–210.

Stranieri, A. and Zeleznikow, J. 1992. SPLIT-UP – Expert system (sic) to determine spousal property distribution on litigation in the Family Court of Australia. In *Proceedings of the Fifth Artificial Intelligence Congress*. Hobart: World Scientific, pp. 51–56.

Terrett, A. 1995. Neural networks – Towards predictive law machines, *International Journal of Law and Information Technology* **3**(1), 95–111.

Thagard, P. 1991. Connectionism and legal inference. *Cardozo Law Review* **13**, 1001–1004.

van Opdorp, G.J. and Walker, R.F. 1990. A neural network approach to open texture. In H.W.K. Kasperen and A. Oskamp (eds), *Amongst Friends in Computers and Law*, Deventer: Kluwer Law and Taxation.

van Opdorp, G.J., Walker, R.F., Schrickx, J.A., Groendijk, C., and van den Berg, P.H. 1991. Networks at work: a connectionist approach to non-deductive legal reasoning. In *Proceedings of the Third International Conference on Artificial Intelligence and Law*. Oxford: ACM Press, pp. 278–287.

Walker, R.F., Oskamp, A., Schrickx, J.A., Opdorp, G.J., and van den Berg, P.H. 1991. PROLEXS: Creating law and order in a heterogeneous domain. *International Journal of Man-Machine Studies* **35**(1), 35–68.

Warner, D.R. 1989. Towards a simple law machine, *Jurimetrics* **29**, 451–467.

Warner, D.R. 1990. The role of neural networks in law machine development. *Rutgers Computer and Technology Law Journal* **16**, 129–144.

Warner, D.R. 1992. A neural network based law machine: Initial steps, *Rutgers Computer and Technology Law Journal* **18**, 51–63.

Warner, D.R. 1993. A neural network-based law machine: The problem of legitimacy, *Law, Computers and Artificial Intelligence* **2**(2), 135–147.

Winston, P.H. 1992. *Artificial Intelligence*. Third edition, Reading, Addison Wesley.

Zurada, J.M. 1992. *Introduction to Artificial Neural Systems*. St Paul: West Publishing.