Santa Clara High Technology Law Journal

Volume 12 | Issue 1 Article 2

January 1996

The Uses and Abuses of Neural Networks in Law

Michael Aikenhead

THE USES AND ABUSES OF NEURAL NETWORKS IN LAW*

Michael Aikenhead†

TABLE OF CONTENTS

† BSc and LLB (Hons), University of Melbourne, Australia, 1994; Candidate, M.Jur., Durham University, England, 1996; currently working as a research assistant in the Centre for Law and Computing, Durham University, England. This research was conducted while attending the University of Melbourne, Australia. The author expresses his thanks to Daniel Hunter for the many helpful comments and suggestions he provided during the preparation of this paper.

I. Introduction

Law has long been an area in which expert system technologies are applied. Numerous legal expert systems, computer systems that perform tasks normally regarded as requiring intelligence, have been created. However, the practical benefit of such systems has been less than predicted. The operation of early systems failed to account for the complexity and subtlety of the law and of legal reasoning.

Researchers in artificial intelligence and law have investigated various proposals to make such systems more realistic. The last decade has seen a resurgence in interest in artificial neural networks (hereinafter neural nets). Neural nets are computer models inspired by biological neural systems in the brain. Researchers believe that by mimicking the underlying structure of the brain they will be better able to mimic the intelligent tasks performed. In the field of law, it is believed that neural nets can overcome some of the limitations associated with existing legal expert systems.

This paper will examine existing and proposed uses of neural nets in the law focusing on the jurisprudential implications and limitations inherent in those proposals. It is divided into six parts. Following this introduction is a technical overview of neural nets, in order to outline their benefits and limitations. This is followed by a discussion of the nature of legal reasoning and the various models proposed to described it. With this background, part four examines various current and proposed uses of neural nets in law. Part five provides a jurisprudential examination of these uses. This paper concludes with some remarks on the uses made of neural nets in the law and the promise they provide for future research into the creation of legal expert systems.

II. Introduction to Neural Nets

A. *Artificial Intelligence*

In the artificial intelligence community, there are several approaches to modelling human intelligence.[1] One approach applicable to the legal domain is the use of symbolic reasoning systems, which are called expert systems. These systems are called symbolic systems because they transform symbols representing things in the real world into other symbols according to explicit rules.[2]

Expert systems have a database of hierarchical rules, variables and constants that they apply to a given problem to try and determine a solution.[3] However, symbolic systems have several limitations, including:

(1) Not all knowledge can be stated symbolically; and
(2) Developing and maintaining the system is time consuming.

The problem of trying to symbolically formalize knowledge can be enormous. It has been found that where there are gray areas to a problem, the resolution of which involves the weighing of a multitude of factors, experts often reach a conclusion and then ex post facto justify it according to their hierarchy of symbolic rules. Rules, then, do not seem to capture all that is involved in expert knowledge.[4]

Secondly, the actual construction and maintenance of the system is complex and time consuming due to a knowledge acquisition bottleneck. The system's creators must explicitly code every rule and predicate manipulated by a symbolic reasoner. The system then has to be debugged to ensure the database is free of errors and operates as predicted. Any changes made to the database, either through changes in or expansion of the knowledge of the system, have to be incorporated through the same time-consuming process. To make these systems

1. The term "modelling human intelligence" is herein used to refer to attempts to model the results achieved by humans when solving problems. How those results are achieved by a human and by a computer may be quite different.

2. Raymond Kurzweil, The Age of Intelligent Machines 16-18 (1990).

3. *See generally* John Zeleznikow & Daniel Hunter, Building Intelligent Legal Information Systems — Representation and Reasoning in Law, ch. 6 (1995) (Computer Law Series No. 13, 1986) (explaining symbolic reasoning using rules and logic).

4. Kurzweil, *supra* note 2.

intelligent requires a great deal of work by a domain expert working in conjunction with a knowledge engineer.[5]

Neural nets adopt an alternative approach to modelling intelligence. In neural nets, the relations between pieces of information do not have to be explicitly specified. Instead, the neural net learns the relationships between the information. For this reason neural nets are subsymbolic reasoners; the system's designers do not have to explicitly state the relationship between pieces of information in the form of symbols. These aspects of neural nets have led to resurgent interest in their use in intelligent computer systems.

B. *Neural Nets*

Neural nets[6] are computer models inspired by the structure of biological neural systems. Biological neural systems are composed of

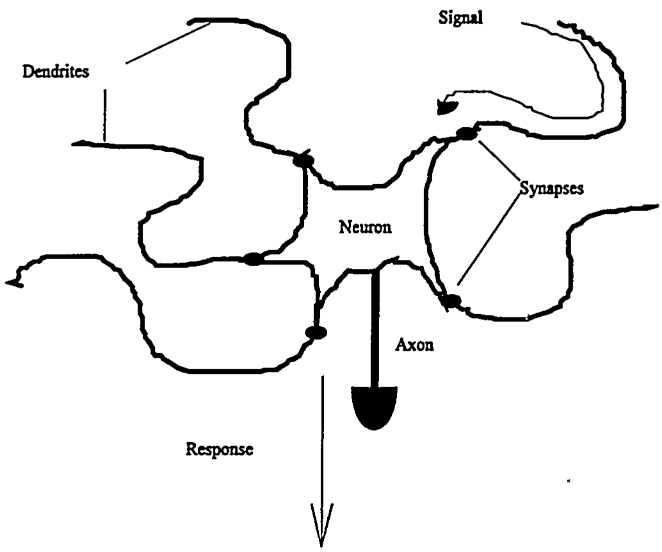


FIGURE 1. STRUCTURE OF A BIOLOGICAL NEURON

5. A domain expert is an expert in the subject matter in which the expert system is sought to be constructed. A knowledge engineer is someone who works with the domain expert to collect that expert's knowledge and assemble it for use in the legal expert system. KURZWEIL, *supra* note 2.

6. *See generally* MAUREEN CAUDILL & CHARLES BUTLER, NATURALLY INTELLIGENT SYSTEMS (1990) [hereinafter NATURALLY INTELLIGENT SYSTEMS]. For a detailed discussion of neural-net concepts and theory, see DAVID RUMELHART ET AL., PARALLEL DISTRIBUTED PROCESSING: EXPLORATIONS IN THE MICROSTRUCTURE OF COGNITION (1986). For a 'hands on' introduction to neural nets, the computer package of Caudill & Butler is useful. *See* MAUREEN CAUDILL & CHARLES BUTLER, UNDERSTANDING NEURAL NETWORKS: COMPUTER EXPLORATIONS (1992).

millions of neurons. Each neuron accepts input from the many thousand other neurons to which it is connected and in turn sends its output to many thousand other neurons. Neurons are connected by axons and by dendrites. A neuron receives signals from other neurons through dendrites and sends its signal to other neurons through its axon.

Where a dendrite connects to another neuron there is a synapse. Synapses are plastic in the sense that the strength of their connection to the neuron can increase or decrease. A strong signal passing through a weak synapse may have the same effect as a weak signal passing through a strong synapse. Synapses can also be inhibitory or excitory. They can either inhibit the activity of the receiving neuron or increase its activity. The inputs that a neuron receives cause it to have some degree of excitation. This level of excitation results in the neuron generating a certain output which it in turn transfers along its axon to the neurons accepting input from it.

Neural nets mimic this structure. Neural nets are composed of neurodes. A neurode is a mathematical model of a biological neuron. Neurodes are connected through synaptic weights to other neurodes, which creates a network. What group of neurodes each neurode accepts input from, what output a neurode generates from its inputs and to which other group of neurodes the output is sent, all determine the way the neural net will behave. One of the major goals in neural net research has been to construct neural nets capable of learning. In biological systems, experiments show that one of the most important effects of learning at the cellular level is the modification of the strength of the synaptic connection between two neurons. Analogously, training a neural net is a matter of modifying the values of the synaptic weights in the system. Unfortunately, training is a complex task, and the method used depends on the architecture of the network in question.

Contrary to the optimistic hopes of early neural net researchers, it is not possible to simply connect many neurodes in a random fashion and hope that they will perform a meaningful task. As in biology, neurodes must be connected in a particular structure.

All neural nets, however, operate as some form of pattern classifier. During its training, the neural net learns to associate a certain pattern presented on its input with a certain pattern on its output. This process is known as pattern association. Further, neural nets have the property that they can generalize their input. This process is known as pattern generalization. Neural nets can learn the characteristics of a general category of objects based on a series of specific examples from that category. This ability to classify patterns is retained even when the neural net is presented with partial patterns. The neural net will infer the general category to which the partial input belongs.[7]

Researchers have experimented with various structures for constructing neural nets ranging from the simplest single neurode to complex hybrid networks. The major drawback in simple networks is that they can only classify linearly separable problems.[8] They cannot be trained to correctly classify every possible collection of patterns.

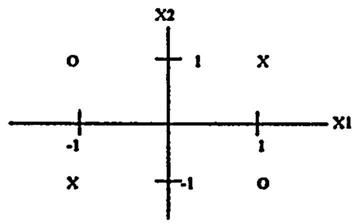
This problem of linear separability can be overcome by using networks of three or more layers. It has been proved that such networks can map any input set to any output set,[9] subject to one limitation: all neural nets, including multilayer neural nets, can only map contradictory input patterns by reaching a compromise between those input patterns. Neural nets cannot take one input pattern and map it to two separate outputs. The consequence of this will be discussed in part five.

Adaptive filter networks are multilayer neural nets and are trained using back-propagation techniques.[10] Although adaptive filter networks must undergo supervised learning,[11] they are perhaps the most common form of neural net used. While much more sophisticated neural nets than adaptive filter networks exist, many are extremely complex and are difficult to implement and tune. For this reason, such networks remain largely at the research stage. Unless

---

7. *See generally* NATURALLY INTELLIGENT SYSTEMS, *supra* note 6, at 3.

8. The classic example of the problem of linear separability is the XOR problem. In the graph below, it is not possible to draw a single straight line that separates all the O's and all the X's, thus they are not linearly separable. *Id.* at 173-74.



9. *Id.* at 174-77 (discussing Kolmogorov's theorem).

10. Back-propagation is a technique whereby the error made by a neural net in classifying a pattern can be progressively reduced, so that it reaches an "acceptable level." *Id.* at 183-96.

11. Supervised learning is a procedure for training a neural net. The neural net is presented with an input pattern and the output that is desired when that input pattern is presented. The neural net learns to associate the input pattern with the output pattern. The learning is supervised because the creator of the neural net must present the two patterns to the network and also oversee that learning is occurring correctly. An obvious requirement of such training is that for every input pattern there must be a known output pattern; this is impossible in some environments, including some legal applications. *Id.*

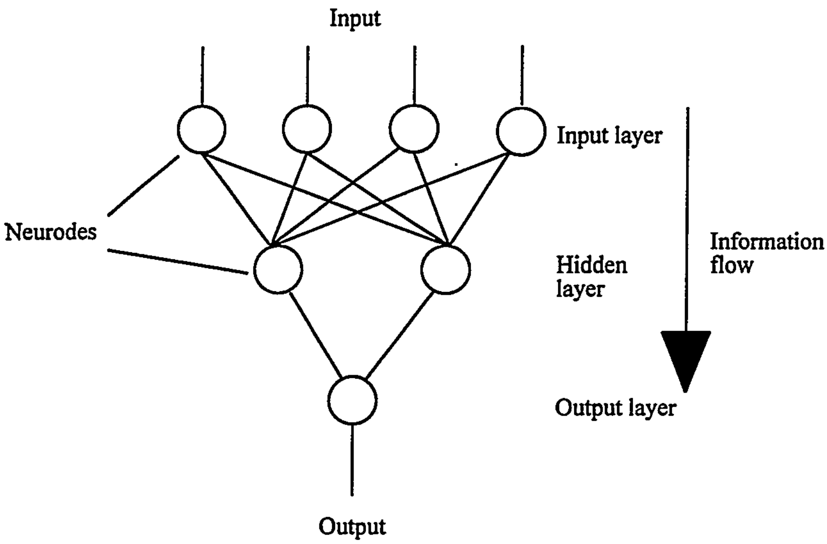


FIGURE 2. GENERAL STRUCTURE OF AN ADAPTIVE FILTER NEURAL NET

otherwise specified, general reference to neural nets in this paper will, thus, concern adaptive filter networks.

C. *Benefits of Neural Nets*

The use of neural networks in the creation of legal expert systems can overcome some of the limitations of symbolic systems. Neural nets can make inferences from incomplete information and classify patterns (both by matching past information and generalizing that past information). This makes them promising candidates for use in various tasks performed by legal expert systems. More importantly, the ability of neural nets to learn may aid in overcoming the knowledge acquisition bottleneck associated with symbolic reasoning systems.

As will be discussed in part four, various current and proposed uses for neural nets in legal expert systems attempt to exploit these properties.

III. INTRODUCTION TO LEGAL REASONING

It is obviously a prerequisite to know what the nature of law is and what the process of legal reasoning involved before incorporating legal knowledge in a computer and making the computer manipulate that knowledge to emulate the legal reasoning process, *i.e.*, the results achieved by lawyers.[12]

This discussion will focus largely on the processes involved in legal reasoning, which are the actual steps undertaken when a lawyer is presented with a problem, decides on the applicable law, applies the law to the problem, and thereby reaches a conclusion. A discussion of the nature of law itself is important to this examination. However, traditional jurisprudential debates such as those between natural lawyers, positivists and realists, over issues such as the validity of the law or the duty to obey the law will not be discussed, as they are peripheral to the present examination.

A. Methods of Reasoning

There are three common methods of human reasoning:

(1) deductive reasoning;
(2) inductive reasoning; and
(3) analogical reasoning.

Detailed expositions of each type of reasoning have been given elsewhere.[13] However, a short explanation is worthwhile.

Deductive reasoning is a strict logical method of reasoning. Deductive arguments take the following general form:

(A) In any case, if p then q;
(B) In the present case p;
(C) Therefore, in the present case, q.

In this form of reasoning, one moves from the application of general rules to specific facts to deduce an outcome. The premises require and justify the conclusion. It is illogical to accept the general rule and the specific instance, but to deny the conclusion.[14] However, the application of the general rule to the specific instance is contingent on that instance being regarded as a member of the general class defined in the rule. In terms of the above example, the "q" referred to in line (C) must be regarded as similar enough to the "q" in line (A) before deductive application of the rule can occur.

Inductive reasoning essentially operates as the reverse of deductive reasoning. Here, one starts with numerous observations and then

---

12. In this paper "lawyer" is used widely to refer to those who are involved in reasoning with and applying the law. Thus it would include judges, solicitors, barristers and legal academics. ZELEZNIKOW & HUNTER, *supra* note 3, at 13.

13. *See, e.g.*, THE ENCYCLOPEDIA OF PHILOSOPHY (Paul Edwards ed., 1967).

14. NEIL MACCORMICK, LEGAL REASONING AND LEGAL THEORY 21-22 (1978).

tries to relate them by creating a rule that can "explain" each observation. For example the following situations are observed:

Facts	Outcome
A B C D E F G	X
A B C D E F	X
A B C D E G K	X
M N O P E	Y

Rules can be stated that explain each outcome;

(A) "If (A B C D E) then X;" and
(B) "If (M N O P E) then Y."

The validity of such rules, though, remains contingent.[15]

Analogical reasoning, in contrast to deductive reasoning and inductive reasoning, is not immediately concerned with the application of rules. Here, one simply says that a certain outcome should result because that outcome has previously occurred in a similar case. This is a manifestation of the formal principle of justice that "like cases are to be treated alike and different cases differently."[16]

Analogical reasoning and inductive reasoning are quite closely related.[17] One only considers that the outcome of two situations should be similar because their facts are similar, by following the rule that like cases should be decided alike. This can itself be seen as the corollary of the more general belief that if two situations have the same outcome, then there must be a general rule that explains them. Thus, in saying that two similar factual situations must have similar outcomes, we are really saying that this would be the result from the application of a hypothetical general rule that would contain both situ-

15. *See* MARTIN GOLDING, LEGAL REASONING 43-44 (1984). It is always possible that a factual situation will arise that has an outcome different to that previously observed, thus invalidating the rule founded on the earlier situations.

16. James Murray, *The Role of Analogy in Legal Reasoning*, 29 UCLA L. REV. 833, 849 (1982).

17. GOLDING, *supra* note 15, at 44 (stating that induction is simply another form of analogy). While they are closely related, there is a difference between analogy and induction, outlined below. *See* discussion *infra* parts V.A.3, V.A.4.

ations.[18] Likewise, it is only possible to suppose that a general rule can explain several situations if one regards those situations as similar to each other.

In this respect, the mental processes in inductive and analogical reasoning are very similar. In both cases, general rules explaining factual situations are assumed to exist. However, only in inductive reasoning does one take the step of trying to explicitly state those rules. More fundamentally, inductive reasoning and analogical reasoning are both inherently dependent on the finding of similarity between situations.

B. *Legal Reasoning*

To what extent does legal reasoning involve each of the above types of reasoning? The answer depends on the nature of the legal system. If the law is a system of rules, the use of induction and analogy will be far more limited than will be the case if law is not totally rule based. Not surprisingly, questions about the nature of law are intertwined with questions about the nature of legal reasoning. As Neil MacCormick states, "A theory of legal reasoning requires and is required by a theory of law."[19]

Two views on the nature of law can be outlined:

> (A) Law is a series of well defined rules of universal application; and
>
> (B) Law is not rule based; legal outcomes are wholly dependent on the views of the parties, lawyers and the judge in a case.[20]

Of course, few if any jurisprudes adhere to these extreme versions of either approach.[21] While it would be fruitless to try to conclusively determine the nature of law, it will be argued that law's true nature does not lie at either of the extremes presented, but incorporates aspects of both positions.

---

18. MacCormick, *supra* note 14, at 163.

19. *Id.* at 229.

20. *See generally* MacCormick, *supra* note 14, at 197, 229; Zeleznikow & Hunter, *supra* note 3, at 53-66. The first of these views is an extreme version of the legal positivism of H.L.A. Hart. *See generally* H.L.A. Hart, The Concept of Law (2d ed. 1994). The second view is an extreme version of the arguments presented by the American legal realists and members of the critical legal studies and post modernist movements. *See generally,* Margaret Davies, Asking the Law Question (1994); Alan Hunt, *The Big Fear: Law Confronts Postmodernism,* 35 McGill L. J. 507 (1990).

21. *See* MacCormick, *supra* note 14, at 197. *But see* Joseph William Singer, *The Player and the Cards: Nihilism and Legal Theory,* 94 Yale L. J. 1 (1984); John Stick, *Can Nihilism be Pragmatic?,* 100 Harv. L. Rev. 332 (1986) (replying to Singer).

Levi has given a useful breakdown of the process of legal reasoning, which he sees occurring in three steps: "(1) Similarity is seen between cases; (2) The rule of law inherent in the first case is announced; and (3) The rule of law is made applicable to the second case."[22] While Levi's description of the legal reasoning process may not capture all that is involved in legal reasoning, it does reveal that perhaps the key step in legal reasoning is the finding of similarity, or difference, between cases and aspects of a case.

In this context, MacCormick and Burton note that the finding of similarity is dependent on the overall purposes that the legal system is trying to achieve.[23] The classification of facts for the purposes of fitting them into the major premise of a deduction and for the purposes of creating analogies and inducing rules, occurs in a whole body of knowledge and theory we use to make sense of the world.[24] When deciding between competing fact classifications, our evaluation inherently involves considerations of the consequences of each classification on our model of the world and in this sense similarities, dissimilarities, classifications, and thus, the meaning and scope of rules are made and not found.[25]

1. Deductive Legal Reasoning

That deduction plays a role in legal reasoning is difficult to deny.[26] Statutes are collections of relatively clearly-stated rules, and thus, the application of statutory law involves a large amount of deductive reasoning.[27] One begins with a statutory requirement, applies it to the facts, and thus, determines the outcome. However, applying the statute to the facts is a complex process.

Firstly, before a deduction can occur, the facts have to be fitted within the language of the statute. This is a nonobvious step, as facts can be logically characterized in several ways.[28] Thus, a logical de-

22. EDWARD LEVI, AN INTRODUCTION TO LEGAL REASONING 1 (1948). *See also* STEVEN BURTON, AN INTRODUCTION TO LAW AND LEGAL REASONING 26-39 (1985) (discussing a similar taxonomy). *But see* Murray, *supra* note 16, at 848-50 (critiquing Levi). For present purposes this criticism is not important.

23. MACCORMICK, *supra* note 14, at 101-108; BURTON, *supra* note 22, at 103.

24. MACCORMICK, *supra* note 14, at 103.

25. *Id.*, ch. 5, ch. 7; Duncan Kennedy, *Freedom and Constraint in Adjudication: A Critical Phenomenology*, 36 J. LEGAL EDUC. 518 (1986).

26. MACCORMICK, *supra* note 14, ch 2; BURTON, *supra* note 22; JULIUS STONE, LEGAL SYSTEM AND LAWYER'S REASONINGS chs. 6-7 (1964); *Cf.* LLOYD OF HAMPSTEAD, LLOYD'S INTRODUCTION TO JURISPRUDENCE 1139, fn. 95 (1985) (for a list of authorities who deny deduction plays a role in legal reasoning).

27. MACCORMICK, *supra* note 14, at 19.

28. BURTON, *supra* note 22, at 44-50; STONE, *supra* note 26, at 55-58.

duction can only occur if the facts are regarded as similar enough to the language of the statute to be classed as covered by the statute. As Edward Levi notes, "The scope of a rule of law, and therefore its meaning depends upon a determination of what facts will be considered similar to those present when the rule was first announced."[29]

Secondly, there is the closely related problem of deciding what meaning to give terms within a statute. For example, section 91(1) of the Crimes Act (1958) (Vic.) states, "A person shall be guilty of an offence, if when not at his place of abode, he has with him any article for use in the course of or in connexion with any burglary . . . ." While classifying an article as within section 91 of the Act may be easy in some cases, this is not always so. What of a tool box? All the items therein could be used during a burglary, yet all could have legitimate uses. Whether an article is for use in the course of a burglary is a matter for debate.

This problem, of determining the meaning of individual phrases in rules, is called the problem of "open texture."[30] Resolving the problem of open texture is inherently dependent on the use of analogy.[31] Thus, even in this, the perhaps most rule-guided area of law, where all the rules are collected and clearly expressed, the purely deductive application of rules is not sufficient to solve all problems.

Similar problems arise when reasoning in the common law. It is often said that there are common law "rules." However, in a strict sense, this cannot be true. The whole of the common law has been created on an individual case by case basis. In a single case, a judge can do no more than pronounce a decision that applies to the facts of the case. It could be argued that the *ratio decidendi* expresses the rule contained in a case.[32] This rule will be binding on all subsequent cases that have the same facts as the original case. However, the binding nature of the *ratio decidendi* (and thus the scope of the rule) is severely limited once it is appreciated that the ratio only applies to the strict facts of the original case. It will only determine the outcome of another case that has exactly the same facts. Strictly, any change in

29. LEVI, *supra* note 22, at 2. Similarly Lloyd notes that it has "long been accepted that a case only binds as to 'like facts.' But what are like facts . . . ." LLOYD, *supra* note 26, at 1116. While given as a discussion of rules in the common law, this is equally applicable to statutory rules.

30. MACCORMICK, *supra* note 14, at 66.

31. *Id.*; LEVI, *supra* note 22; BURTON, *supra* note 22.

32. RUPERT CROSS & J. HARRIS, PRECEDENT IN ENGLISH LAW (1991).

the facts results in a new situation, the outcome of which is not determined by the *ratio decidendi*.[33]

The belief that there are common law rules arises because, even though the *ratio* of one case may not be binding in a later case, if the latter case has very similar facts to the original case, the *ratio* is nevertheless felt to be highly persuasive.[34] Thus, the second case is decided similarly to the first. As this process continues, a large body of cases builds up, all of which have similar facts and similar outcomes. Seeing this collection of cases, it is not unreasonable to assume that the original case laid down a general rule, which dictated the results in all the latter cases.[35] In this way, the common law appears to create rules that can later be applied deductively.[36] Even in such usage though, these common law rules experience the same problems as statutory rules.

2. Inductive and Analogical Legal Reasoning

The process of induction will often be used in framing the *ratio* of leading cases and the construction of novel legal arguments. Before a leading case, there often exists numerous cases with vaguely similar facts and similar outcomes. However, each has been decided in a relatively individual way. When a leading case is decided, the judges will look at the previous cases and surmise that, since they have similar facts and similar outcomes, there must be a general rule or unifying doctrine that explains all the cases. In this way, the general rule-like pronouncements contained in the *ratio* of the leading case will have been induced from the previous cases.[37] The same process occurs when counsel advances a new unifying rule in argument. As with reasoning by analogy, the use of inductive reasoning in the law can be seen as an inherent consequence of the requirement for coherence within the legal system as enunciated by MacCormick.[38]

---

33. MACCORMICK, *supra* note 14, at 219-24; STONE, *supra* note 26, at 267-74. Indeed, Stone regards the multitude of *ratios* that exist in a decision as requiring extreme skepticism about the ability of computers ever to reason with cases. *Id.* at 37-38; *Cf.* CROSS & HARRIS, *supra* note 32.

34. This results from the need for reality and coherence in the legal system. MACCORMICK, *supra* note 14.

35. *Id.* at 216-218.

36. In truth though, common law rules only serve to hide the cases underlying the supposed rule and to mask the reaching of a decision by analogy. BURTON, *supra* note 22, at 60; LEVI, *supra* note 22, at 8-9.

37. CROSS & HARRIS, *supra* note 32, at 191-192. Levi states that thinking of case-law reasoning as inductive is erroneous. However, he agrees that case law concepts can be created out of particular instances, since there is movement from the particular to the general. LEVI, *supra* note 22, at 27.

38. MACCORMICK, *supra* note 14, ch. 7.

The use of analogical reasoning in law has been widely studied.[39] As with deductive legal reasoning and inductive legal reasoning, such descriptions emphasize the necessity of finding similarities between cases before any analogy can be constructed.[40]

Assume that the following cases have the following factors:

Case 1: A B C D;
Case 2: A B C E;
Case 3: A B C F; and
Case 4: A B C G.

Further, assume that Case 1 and Case 2 are regarded as analogous. From this it can be implied that factors D and E are similar. Again, assume that Case 1 and Case 3 are not regarded as analogous, implying that factors D and F were not similar. How is Case 4 to be classified? This depends on whether factor G is regarded as more similar to factor D or more similar to factor F.

A consequence of the importance of the finding of similarity to the process of legal reasoning is that extreme versions of legal positivism do not seem supportable. Since deductive reasoning is by itself insufficient to explain legal reasoning, law must be composed of more than purely rules. Nor, however, can it be accepted that legal reasoning is totally subjective;[41] legal rules provide a paradigm that guides legal thought.[42]

This view of legal reasoning, as a process inherently dependent on the finding of similarity between situations and on our world theories, has consequences for the use of neural nets in legal expert systems.

IV. Current and Proposed Uses of Neural Nets in the Law

Mirroring resurging interest in the general artificial intelligence community, the use of neural nets in law has recently received growing interest. Neural nets have been used and have been proposed to be used in the law in two broad manners:

39. *Id.*; Levi, *supra* note 22; Burton, *supra* note 22; Stone, *supra* note 26; Murray, *supra* note 16; James Gordley, *Legal Reasoning: An Introduction*, 72 Cal. L. Rev. 139 (1984); Cass Sunstein, *On Analogical Reasoning*, 106 Harv. L. Rev. 741 (1993).

40. *E.g.*, Stone, *supra* note 26, at 283.

41. Steven Burton, *Reaffirming Legal Reasoning: The Challenge from the Left*, 36 J. Legal Educ. 358 (1986). *Compare* Singer, *supra* note 21 *with* Stick, *supra* note 21.

42. K. Hamilton, *Prolegomenon to Myth and Fiction in Legal Reasoning, Common Law Adjudication and Critical Legal Studies*, 35 Wayne L. Rev. 1449 (1989).

(1) as and within inference engines[43] in legal expert systems; and
(2) in legal information retrieval systems.

This part will discuss and explain each of these proposed uses. The following part will discuss some of the jurisprudential implications arising from these proposed uses.

A. *Neural Nets As and Within Inference Engines*

Before examining the current and proposed uses of neural nets in the law, it is beneficial to have an understanding of more traditional techniques for computerizing legal knowledge.

As Stephen I. Gallant explains, there is no definitive definition of "what constitutes an expert system."[44] However, for the purposes of this discussion, the following loose definition will be adopted: a legal expert system is a computer program capable of performing tasks usually performed by a lawyer "at the standard of (and sometimes even at a higher standard than) human experts in given fields."[45] In this respect, expert systems and information retrieval systems are very similar: both require aspects of intelligence. However, only in an expert system does the system try to reason with the law.

1. Traditional Legal Expert System Inference Mechanisms

It has been widely noted in the artificial intelligence and law research community that the dominant reasoning paradigm used in legal expert systems is that of symbolic reasoning.[46] Both production rule expert systems and symbolic case based reasoners[47] adopt this approach. Production rule expert systems seek to encode law in the form of rules of logic. Symbolic case based reasoners encode aspects of cases, such as the factual attributes, which then undergo transformations and are reasoned with according to explicit rules.[48]

---

43. An inference engine is a part of an expert system that is "a system for applying the rules [of the system's database] to the knowledge base to make decisions." KURZWEIL, *supra* note 2, at 293.

44. STEPHEN I. GALLANT, NEURAL NETWORK LEARNING AND EXPERT SYSTEMS 255-261 (1993).

45. ZELEZNIKOW & HUNTER, *supra* note 3, at 68. The authors note that the actual task that a legal expert system will perform varies markedly according to its intended user. *Id.*

46. *E.g.*, Kenneth Lambert & Mark Grunewald, *Legal Theory and Case-Based Reasoners: The Importance of Context and the Process of Focusing*, THIRD INT'L CONF. ON ARTIFICIAL INTELLIGENCE & L. 191 (1991).

47. Case based reasoners are expert systems that try to reason using a corpus of cases rather than explicit rules. ZELEZNIKOW & HUNTER, *supra* note 3, at 182.

48. John Zeleznikow et al., *The IKBALS Project: Multi-Modal Reasoning in Legal Knowledge Based Systems*, 2 ARTIFICIAL INTELLIGENCE & L. 169, 171-72 (1993).

2. Problems with Traditional Inference Mechanisms

A major jurisprudential problem with symbolic reasoning systems is that they depend upon legal knowledge being composed of explicitly stated rules. As argued in the previous part, however, law is composed of more than just rules.

Further, the developers of symbolic reasoning systems assume that deductive reasoning is the only mode of reasoning applied in the law. Although analogical reasoning is emulated in symbolic case-based reasoners, what is actually implemented in such systems is only a crude simulation of human analogical reasoning. Symbolic case-based reasoners rely on explicit rules of how cases and case attributes can be manipulated. It is assumed that, through the deductive application of these rules, analogical reasoning will emerge. As a complete model of analogical reasoning, this is a dubious assumption.

Thirdly, symbolic reasoners experience difficulties in resolving conflicts between rules in their rule databases.[49] Such conflicts can only be resolved with metarules. Again, this assumes that the law is a deepening spiral of rules, which is jurisprudentially suspect.[50]

Finally, there are practical problems in the creation of symbolic reasoners. Such systems experience a knowledge acquisition bottleneck.[51] Any changes in the relevant law will require a modification of the database, which then needs to be debugged — a time-consuming process.

The use of neural nets in legal expert systems has been proposed as a means to overcome these problems. These proposals will be examined below.

3. Proposed Uses for Neural Nets in Legal Expert System Inference Mechanisms

(a) Reasoning with Cases

It was argued in part three that the law cannot be regarded purely as a system of rules; analogical reasoning from past cases is extremely important. Neural nets may find application in systems that reason

---

49. For example, if the database contained the two rules "As between two innocents he who caused the damage should pay" and "No liability without fault" it is unclear how these rules are both to be applied in a no-fault accident. The resolution of this conflict must be resolved by reference to other tests.

50. DAVIES, *supra* note 20, ch. 7 (demonstrating how both Hart's concept of a rule of recognition and Kelsen's concept of a grundnorm would necessarily import extra-legal assumptions into the legal system).

51. *See supra* part II (discussing the problem of knowledge acquisition bottleneck in symbolic reasoners).

with cases. David Warner has stated that a large benefit of neural nets is their ability to classify patterns and so imitate the analogical reasoning process, thereby resolving issues of open texture.[52] However, before this claim can be sustained, the precise nature of the analogical reasoning process needs to be investigated.

i. Analogy

As outlined in the previous part, an important aspect of analogical reasoning is the classification of patterns into similar groups; only things that are similar can be used in an analogical argument. Neural nets are inherently good at pattern classification,[53] which makes them seemingly promising candidates for emulating the analogical reasoning process.

The use of neural nets to mimic this aspect of analogical reasoning has been investigated by Hobson and Slee. They have produced a neural network "index" of the Theft Act 1968 (England).[54] In this index, a factual situation is analyzed by the researchers for the presence or absence of various concepts, the concepts being specified by the wording of the Act. The presence or absence of each concept results in a matrix that is then used as the input to their neural net. The verdict on whether or not the factual situation constitutes theft within the meaning of the terms of the Act is used as the desired output for the neural net.

Using this material, Hobson and Slee claim a neural net can be trained to classify cases covered by the Act.[55] During training, the neural net groups the cases used to train it into general groups.

Once trained, new cases can be presented to the neural net. In reaching a verdict on a new case, the neural net classifies the case into one of the general groups created during training. In so classifying a case, the neural net appears to mimic analogical reasoning; similar cases result in the same verdict.

---

52. *See* David R. Warner, Jr., *A Neural Network Based Law Machine: Initial Steps*, 18 RUTGERS COMPUTER & TECH. L.J. 51, 51-54 (1992); David R. Warner, Jr., *The Role of Neural Networks in the Law Machine Development*, 16 RUTGERS COMPUTER & TECH. L.J. 129, 139 (1990) (discussing how a neural network can be used as a model to resolve the "open-texture" problem in the legal reasoning process).

53. All neural nets operate as some form of pattern classifier. They learn to associate certain general input patterns with certain general output patterns. *See supra* part II.

54. John Hobson & David Slee, *Indexing the Theft Act 1968 for Case Based Reasoning [CBR] and Artificial Neural Networks [ANNs]*, FOURTH INT'L CONF. ON ARTIFICIAL INTELLIGENCE & L., COMPUTERS & ARTIFICIAL INTELLIGENCE (1994).

55. Success in this respect must be understood to mean performing a classification, according to the index points chosen by the creators, which is the same as that which the creators would arrive at using those same index points.

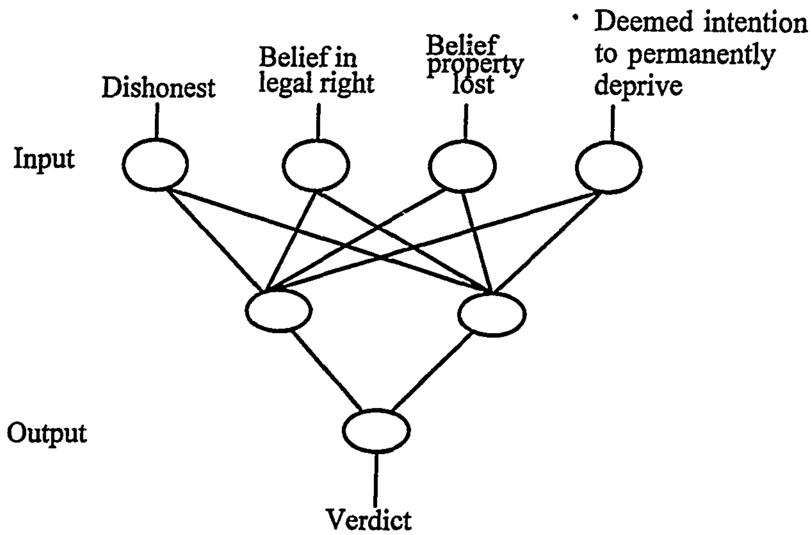


FIGURE 3. NEURAL NET INDEX OF THEFT ACT 1968 (ENGLAND)

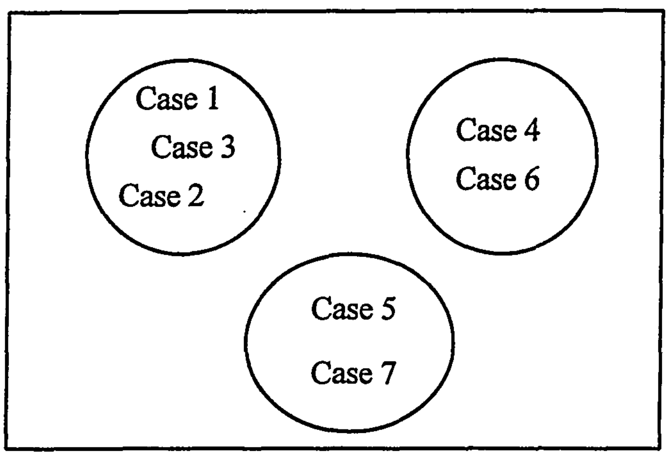


FIGURE 4. NEURAL NETS GROUP SIMILAR CASES TOGETHER

Similar work has been performed by Trevor Bench-Capon, who has created a neural net based on a hypothetical statute.[56] Bench-Capon's investigation is of further interest in that it demonstrates that a neural net can successfully perform classifications even when presented with a lot of "noise" (inputs that are not relevant to the classification). Thus, contrary to what other commentators have said,[57] neural nets have the potential to operate successfully even when the factors affecting the classification are not known.[58]

In contrast to the above two approaches, which essentially try to model whole areas of law using neural nets, Walker et al. (the "VUA team") simply use neural nets within a more conventional case based reasoning system.[59] The VUA team have created PROLEXS, a hybrid legal expert system, which relies on more than one model of legal reasoning. Early versions of the system operated by having a stored database of cases, each case being stored as a set of conditions each with an associated fixed weight, along with a case threshold.[60] When a case was to be applied analogically, the weights on conditions present in the current fact situation were summed and then compared to the case threshold of the past case to determine whether the current situation was analogous to the stored case.[61] In the first implementations of PROLEXS, the condition weights and the threshold values had to be assigned by the domain expert. However, the VUA team noted that weight and threshold assignment is a difficult task for a human domain expert.[62] Consequently, the latest version of PROLEXS dispenses with the case database within the case based reasoning subsystem. Instead, a multilayer neural net is trained using the conditions as the inputs and the applicability or nonapplicability of the open texture term as the output to the neural net.[63] The neural net learns the condition weights and case thresholds during its training. This is essentially the same approach as taken by Hobson and Slee, and

56. Trevor Bench-Capon, *Neural Networks and Open Texture*, FOURTH INT'L CONF. ON ARTIFICIAL INTELLIGENCE & L. 292 (1993).

57. G. van Opdorp, et al., *Networks at Work: a Connectionist Approach to Non-deductive Legal Reasoning*, THIRD INT'L CONF. ON ARTIFICIAL INTELLIGENCE & L. 278 (1991).

58. While the ability of a neural net to classify patterns even in the presence of noise is notable, as will be discussed in the next part, defining input as "noise" is dependent on a preexisting theory of the domain. This may be problematic. Bench-Capon, *supra* note 56, at 296.

59. R. F. Walker et al., *PROLEXS: Creating Law and Order in a Heterogenous Domain*, INTER'L J. MAN-MACHINES STUDIES 35 (1991); *see also* van Opdorp, *supra* note 57.

60. Walker, *supra* note 59, at 55-56.

61. *Id.* at 56.

62. *Id.* at 56-57. *See also* van Opdorp, *supra* note 57, at 280-81.

63. Van Opdorp, *supra* note 57, at 281-84.

Bench-Capon. It is claimed that this system can provide more discerning weights than a human expert.[64]

ii. Open Texture

Bench-Capon states that in creating his neural net the ability of neural nets to perform classification in domains involving open texture has been demonstrated.[65] This claim must be questioned. Bench-Capon did not use a neural net to model solely open textured aspects of the domain, but the whole domain itself. This can be regarded as an ironic response to Sergot's attempt to model law purely using rules[66] and the criticism and comment that this created.[67]

Similarly, when creating their neural net, Hobson and Slee treat issues such as whether an action was dishonest (which is an open textured issue) as simply any other index point.[68] It is up to the creators (and presumably later users) of the neural net to decide whether these concepts are present before input is given to the neural net.

The work by Hobson and Slee and Bench-Capon, thus, should not be viewed strictly as a demonstration of the ability of neural nets to resolve open texture, but simply as a demonstration of the ability of neural nets to classify legal cases into desired legal categories.

At this point, the ability of neural nets to resolve issues of open texture may be doubted; this is not an area that has yet been directly investigated. The work, however, has a prima facie appeal: in using neural nets to classify cases, it appears that analogical reasoning is being emulated and it is by analogical reasoning that open texture may be resolved. In the next part the uses of neural nets in computerized analogical reasoning systems will be discussed in more detail, thereby giving more credence to the claim that neural nets can aid in the resolution of open texture.

---

64. *Id.* at 280-81.

65. Bench-Capon, *supra* note 56, at 297.

66. *See* M. J. Sergot, et al., *The British Nationality Act as a Logic Program*, 29 COMM. OF THE ACM 370 (1986) (describing how the British Nationality Act was translated into a form of logic for expressing and applying legislation).

67. Refer to the debate conducted in the following articles: Robert Moles, *Logic Programming—An Assessment of its Potential for Artificial Intelligence Applications in Law*, 2 J.L. & INFO. SCI. 137 (1991); John Zeleznikow & Daniel Hunter, *Rationales for the Continued Development of Legal Expert Systems*, 3 J.L. & INFO. SCI. 94 (1992); Robert Moles & Surendra Dayal, *There is More to Life than Logic*, 3 J.L. & INFO. SCI. 188 (1993); Daniel Hunter et al., *There is Less to This Argument than Meets the Eye*, 4 J.L. & INFO. SCI. 46 (1993).

68. It is unclear what use Hobson and Slee intend for their index. If it is truly meant to be used as an index of cases then their treatment of open textured issues is less questionable than if they intend it to be used within a legal expert system.

(b) Dealing with Conflicting Rules

A number of researchers have proposed the use of neural nets to overcome the difficulty of reasoning with conflicting rules of law. David Warner claims that neural nets can inherently model the legal reasoning process and that neural nets can model a legal system in which rules conflict by giving those rules weights.[69] However, no details are given as to how this is to be achieved.

Lothar Philipps agrees that the ability to model conflicting rules is a benefit of neural nets and has created a neural net designed to investigate this concept.[70] Philipps claims that his neural net can mimic the results that German courts reach when they assign liabilities in automobile accidents and, specifically, that his neural net can mimic the process that occurs when a court is presented with contradictory cases. Contradiction is dealt with by reaching a compromise solution to the conflict.[71] However, while the ability of neural nets to reach compromise solutions is important, as will be discussed later, it may not always be jurisprudentially desirable.

In contrast to the approaches of Warner and Philipps, who model conflicting law on a rule by rule basis through the use of compromise, Paul Thagard has developed a theory of explanatory coherence that he says can "choose" between competing hypothesis. Thagard's system, ECHO, models competing theories using a neural net.[72] ECHO chooses between conflicting groups of rules[73] and does not deal with the conflict through compromise, but accepts or rejects one of the hypothesis.[74]

The notion of competing theories should be familiar to lawyers and legal theorists. In any conflict, there are always at least two competing theories of the law and the facts: that presented by the plaintiff

69. David R. Warner, Jr., *The Role of Neural Networks in the Law Machine Development*, 16 RUTGERS COMPUTER & TECH. L. J. 129, 135-38 (1990) [hereinafter *Role of Neural Networks*]. The claim that neural nets can inherently model the process of legal reasoning will be critically discussed *infra*, part V.

70. Lothar Philipps, *Distribution of Damages in Car Accidents Through the Use of Neural Networks*, 13 CARDOZO L. REV. 987 (1991).

71. *Id.* at 989.

72. Paul Thagard, *Explanatory Coherence*, 12 BEHAV. & BRAIN SCI. 435, 439 (1989) [hereinafter Thagard, *Explanatory Coherence*]; Paul Thagard, *Connectionism and Legal Inference*, 13 CARDOZO L. REV. 1001 (1991) [hereinafter Thagard, *Connectionism and Legal Inference*]). *See also* GALLANT, *supra* note 44, chs. 14-15 (explaining how to create inference networks generally using neural nets.)

73. Though the system could logically be used to resolve conflicts between single rules.

74. *See* Thagard, *Explanatory Coherence*, *supra* note 72, at 435. ECHO requires competing hypothesis to be given to the system, along with evidence, details of how each hypothesis explains the evidence and details of how the propositions are contradicted by the evidence. ECHO then determines which hypothesis best coheres with the evidence. *Id.*

and that presented by the defendant. Prima facie, ECHO seems to provide a possible way to model this conflict in a legal expert system. Thagard has applied this theory in a simplistic manner to two legal situations. However, the possibilities in this approach remain largely unexplored.[75]

(c) Machine Learning

Along with the above problems with knowledge representation and manipulation in symbolic reasoning systems, a further problem with such systems is the expense in developing and maintaining the knowledge base of the system.[76] Neural nets, in contrast, learn their knowledge, which provides a further attraction to their use in legal expert systems.

If a neural net is used to store cases, as in some of the work by Hobson and Slee and in later versions of PROLEXS, then adding new cases to the neural net's knowledge base is simply a matter of presenting those cases to the neural net while it is in its learning mode. The neural net automatically incorporates the cases into its knowledge base through modifying its internode weights. Cases still have to be described in terms amenable to use in the neural network. However, later rerationalizations of those cases[77] only require the net to be retrained, rather than requiring the complete reentry of a newly structured case database.

An additional way to use neural nets to overcome the knowledge acquisition bottleneck is through their use in rule induction systems.[78] Such systems attempt to model the process of legal induction by examining numerous cases and attempting to find relations between factors in those cases that can be explained with a rule.

Although several researchers have proposed the extraction of rules from neural nets to enhance their explanation facilities,[79] neural nets have also been used in an attempt to extract rules from a corpus of cases. In their work on the MAIRILOG project, Laurent Bochereau

75. *See* Stick, *supra* note 21, at 363 (noting that many contemporary theories of law are based upon coherence theories of truth). Thagard's ECHO could be useful in investigating such theories.

76. Richard Susskind, *Expert Systems in Law: A Jurisprudential Approach to Artificial Intelligence and Legal Reasoning*, 49 MOD. L. REV. 168, 184 (1986).

77. *See, e.g.*, Robert Birmingham, *A Study After Cardozo: De Cicco v Schweizer, Non-cooperative Games, and Neural Computing*, 47 U. MIAMI L. REV. 121 (1992). A rerationalisation of a case occurs when a later case rationalises the decision in an earlier case on grounds that are different from those stated in the judgement of the earlier case.

78. *See* discussion *infra* part V.A.1.

79. For example, Bench-Capon, *supra* note 56; van Opdorp et al., *supra* note 57, at 285; GALLANT, *supra* note 44, at 315. *See* discussion *infra* part V.A.1.

et al. propose methods by which logical rules can be extracted from trained neural networks.[80] The rules extracted from the neural net could then be incorporated into a symbolic reasoning system. Again, the ability of neural nets to learn new cases is exploited.

Symbolic methods also exist to extract rules from a corpus of cases, and, as both the use of neural nets and symbolic methods rely on a statistical analysis of data, attempts to use neural nets to induce rules share the limitations of these symbolic automatic rule induction systems.[81] However, the potential for neural nets to incorporate more flexible notions of analogy[82] than those currently used in other induction systems, may overcome some of the limitations currently inherent in automatic rule induction.

B. *Use of Neural Nets in Legal Information Retrieval Systems*

In contrast to the above neural net applications, which try to reason with the law, neural nets have also been used in systems that solely try to retrieve information. Computers have long been used to automate legal information retrieval. However, all have suffered limitations.[83] Most notably, they are brittle in that they rely on keyword searches.[84] Systems employing neural nets remove some of these limitations.

One of the most interesting legal information retrieval systems using neural nets is SCALIR.[85] Created by Daniel Rose and Richard Belew, SCALIR is a combination of a neural net embodying subsymbolic information integrated with a semantic network[86] embodied in a neural net. SCALIR can perform impressive document retrieval

80. *See* Laurent Bochereau et al., *Extracting Legal Knowledge by Means of a Multilayer Neural Network Application to Municipal Jurisprudence,* THIRD INT'L CONF. ON ARTIFICIAL INTELLIGENCE & L. 288 (1991). *See also* Bench-Capon, *supra* note 56, at 296. *See generally* GALLANT, *supra* note 44, ch. 17 (presenting a detailed discussion of the extraction of rules from neural nets).

81. ZELEZNIKOW & HUNTER, *supra* note 3, at 272.

82. *See* discussion *infra* part V.A.3.a.

83. For a comprehensive discussion of legal information retrieval systems and methods and their associated limitations, see generally ZELEZNIKOW & HUNTER, *supra* note 3, at 29-52.

84. For example, if a document is indexed on the term "solicitor" then searching for "lawyer" will not retrieve it, even though it may be relevant. While this problem can be reduced using a search on all synonyms this does not guarantee all relevant documents will be retrieved. *Id.* at 34.

85. Daniel E. Rose & Richard K. Belew, *Legal Information Retrieval: A Hybrid Approach,* SECOND INT'L CONF. ON ARTIFICIAL INTELLIGENCE & L. 138 (1989) [hereinafter Rose & Belew, *Legal Information Retrieval*]; Daniel E. Rose & Richard K. Belew, *A Connectionist and Symbolic Hybrid for Improving Legal Research,* 35 INT'L J. MAN-MACHINE STUDIES 1 (1991) [hereinafter Rose & Belew, *Connectionist and Symbolic Hybrid*].

86. For a discussion of semantic networks and knowledge representation see generally ZELEZNIKOW & HUNTER, *supra* note 3, ch 7.

operations. For example, Rose and Belew show how SCALIR can retrieve a copyright case using a term that does not occur in that case. SCALIR is able to overcome superficial differences in the topic through its network of term associations. This is more than a simple synonym search, though, as the system could also perform such a retrieval based on the fact that two cases have cited a common case.

An interesting aspect of SCALIR is that one of the neural nets is used not to store conceptual features of the domain, but rather microfeatures, such as the fact that a particular word is used in a case or the fact that two cases are often retrieved together.[87] This contrasts with the use made of neural nets when reasoning with cases, where the inputs and outputs were all at the conceptual level.

A further notable feature of SCALIR is its ability to learn from interaction with its users.[88] The system modifies the weights on links within the networks, depending on the type of searches performed by users. Thus, Rose and Belew claim that over time the system can adapt to new terminologies and "the changing importance of cases and statutes."[89]

SCALIR is an impressive advance over prior information retrieval systems and, as Rose and Belew indicate, the benefits provided by SCALIR should not be regarded as lying solely in the field of legal information retrieval.[90] It will be argued in the following part that the combination of subsymbolic and semantic information contained in SCALIR's neural networks is a powerful and flexible method of emulating the finding of similarity required in legal reasoning. Thus, the techniques embodied in SCALIR could find wider application as part of a legal expert system.

The above discussion has provided an introduction to the uses that neural nets have been put in the law. However, there a number of problems and concerns associated with these uses of neural nets. The following part will discuss several more theoretical proposals for the use of neural nets in law and then proceed to examine problems and concerns inherent in these proposals and the uses discussed above.

87. Rose & Belew, *Legal Information Retrieval*, *supra* note 85, at 141.

88. Rose & Belew, *Connectionist and Symbolic Hybrid*, *supra* note 85, at 20-22.

89. *Id.* at 22.

90. *Id.* at 30.

V. Jurisprudential and Technical Concerns in the Use of Neural Nets in Law

A. *Jurisprudential Concerns about Proposed Uses*

As stated in part three, the nature of law and the nature of legal reasoning are two issues that are inherently intertwined. There is an oft-made claim that legal expert systems will provide information about the nature of law and the process of legal reasoning.[91] It is hoped that the use of neural nets in legal expert systems will aid in this. However, for jurisprudence to gain from the creation of legal expert systems and specifically from the use of neural nets, lawyers and legal theorists must be confident that the use of those neural nets rests on a solid jurisprudential basis.

This part will commence with a discussion of two claims about the nature of law that neural nets are said to be suited to modelling. Following this discussion are several jurisprudential observations specifically concerning current and proposed uses of neural nets in the law. Finally, a discussion of how neural nets can offer a new metaphor of law will be presented.

1. Law as a Parallel Process

Amongst those researchers who advocate the use of neural nets in the law, perhaps the most wide-ranging and controversial position is that taken by Warner.[92] Warner is of the view that legal reasoning is an inherently parallel process[93] and that "when we attempt to model the legal reasoning process, we must use a device capable of emulating the parallel problem-solving process. To this end, normal digital computational devices are inadequate."[94] It is claimed that neural networks will overcome this problem due to the inherently parallel nature of their operation.[95] If taken to its full extreme, Warner's view of the legal reasoning process as inherently parallel has potentially fatal consequences for traditional symbolic systems. However, apart from such vague and dubious observations about the nature of legal reasoning, the full implications of this idea are not explored.

---

91. Susskind, *supra* note 76.

92. *See generally* David Warner, *Toward a Simple Law Machine*, 29 Jurimetrics J. 451 (1989) [hereinafter Warner, *Simple Law Machine*]; David Warner, *A Neural Network Based Law Machine: Initial Steps*, 18 Rutgers Computer & Tech. L.J. 51 (1992) [hereinafter Warner, *Initial Steps*]; Warner, *Role of Neural Networks*, *supra* note 69.

93. Warner, *Simple Law Machine*, *supra* note 92, at 461-64; Warner, *Role of Neural Networks*, *supra* note 69, at 131-32.

94. Warner, *Initial Steps*, *supra* note 92, at 53.

95. *Id.* at 53-54.

That the legal reasoning process is an inherently parallel process is highly contentious. It seems acceptable to say, as Warner claims, that when problems are solved, the solution of "unit problems" will impose "a state change on the problem domain rendering invalid all unit solutions previously achieved and changing the environment for all unit solutions yet to be achieved."[96]

However, this is not a description of a parallel process. This simply notes that the answer to one question may change which questions are subsequently asked. While this undoubtedly occurs in human reasoning, the contingent nature of questions is quite easily represented in a tree diagram.[97] Such tree diagrams form the basis of all rule based expert systems. Systems such as PROLEXS[98] display this "parallel" problem solving capability by modifying subsequently asked questions according to intermediate answers. This belief that neural nets can solve all the problems that currently beset symbolic legal expert systems is, perhaps unconsciously, echoed by Bench-Capon. He has attempted to model what, prima facie, appears a rule based area of law with a neural net.[99]

If it is accepted that some legal reasoning occurs in parallel, it still does not mean that all legal reasoning does. It is not in every legal question that, as Levi would say, the application of the rule changes the rule itself. Thus, Warner's vision of the necessity of using neural nets to model the supposedly parallel nature of the legal reasoning process cannot be supported.

2. Open Texture as Randomness

A further contention made by Warner is his equating of the concept of open texture with the idea of randomness.[100] If this view of open texture is correct, then little hope can be held for the ability of lawyers and legal theorists, let alone neural nets or any other legal expert system, to resolve issues of open texture. After equating open texture and randomness, it is surprising that Warner then claims that the use of neural nets can overcome this problem.[101]

---

96. Warner, *Role of Neural Networks*, *supra* note 69, at 132.

97. *See* ZELEZNIKOW AND HUNTER, *supra* note 3, at 118-25. *See generally* ALAN TYREE, EXPERT SYSTEMS IN LAW (1989) (discussing the use of logic and tree diagrams in representing laws).

98. Walker, *supra* note 59.

99. Bench-Capon, *supra* note 56.

100. Warner, *Role of Neural Networks*, *supra* note 69, at 138-39.

101. *Id.*

Perhaps Warner's choice of the term "randomness" was ill-advised. Warner cites and accepts the work of Gardner[102] in his argument for the benefits that neural nets could provide.[103] However, Gardner suggests that open texture can be dealt with through the use of heuristics,[104] a claim Warner accepts.[105] Logically, though, if heuristics exist in a domain, then that domain cannot be regarded as truly random. Warner is correct however in viewing open texture as an "indeterminacy."[106]

As argued in part two, the resolution of open texture does not occur unconstrained, but proceeds through a process of analogy from past cases. During this reasoning process, the factors that can be taken into account and the manner in which they can be used are both constrained.[107] However, the work of critical legal scholars does show that extra-legal factors may influence which factors will be emphasized (or even considered) in a decision. This could make strictly legal examinations of those decisions conclude that they were random.[108]

Thus, the claim that open texture involves randomness does have merit in that it highlights the unpredictability of solutions. Even when all the past cases have been rationalized, the possibility remains that this rationalization will be destroyed if the present case is decided in a novel manner; in neural net terms, with the addition of a new input factor.[109] However, even this possibility is constrained by the need for coherence within the legal system and by the need for the distinction made to be justifiable.[110] So while the resolution of open texture may be difficult and sometimes unpredictable, it is not random.

This resolution of problems involving open texture is inherently intertwined with the nature of analogical reasoning, a subject that will be discussed next.

---

102. ANNE VON DER LIETH GARDNER, AN ARTIFICIAL INTELLIGENCE APPROACH TO LEGAL REASONING (1987).

103. Warner, *Role of Neural Networks*, *supra* note 69, at 139.

104. GARDNER, *supra* note 102, at 41-43 (discussing heuristics as "rules of thumb" used by experts in a field).

105. Warner, *Role of Neural Networks*, *supra* note 69, at 139.

106. *Id.*

107. Kennedy, *supra* note 25; Jorgen Karpf, *Inductive Modelling in Law: Example Based Expert Systems in Administrative Law*, PROC. OF THE THIRD INT'L CONF. ON ARTIFICIAL INTELLIGENCE & L. 297, 300 (1991) (noting that only certain combinations of factors are legal combinations).

108. *E.g.*, Kennedy, *supra* note 25.

109. Birmingham, *supra* note 77, at 132-34.

110. MACCORMICK, *supra* note 14.

3. Analogy and Explanation

Every legal expert system necessarily embodies a jurisprudential theory.[111] The use of neural nets in legal expert systems will affect the nature of that jurisprudential theory.

(a) Analogy

As noted in part two, a key step in analogical reasoning is the finding of similarity, or difference, between cases and aspects of a case. Of course, legal analogical reasoning is much more than the mere finding of similarity between cases. Once two cases are found to be similar, there are limitations on the way the cases can be applied.[112] However, when can two things be regarded as the same or different?

i. Theories of Similarity

Mital and Johnson note that there are no entirely satisfactory theories of what constitutes similarity.[113] According to the ruleless theory, there are no general principles applied in a finding of similarity, people know it when they see it.[114] An alternative, that similarity is found solely by calculating the number of shared attributes that are present in a situation cannot be accepted.[115] If the ruleless theory of similarity is accepted, then little hope can be held for any formalization of the process of finding similarity. However, if it is accepted that some guidelines are followed, it must be appreciated that it is not merely the number of attributes that are shared, but also their relevance.[116]

In this context, Celeste Tito has said that two things are necessary for computers to understand "similarity":

(1) They must understand the analogue meaning of words; and,

111. Susskind, *supra* note 76, at 183.

112. *See* MacCormick, *supra* note 14; Levi, *supra* note 22; Burton, *supra* note 22; Gordley, *supra* note 39; Sunstein, *supra* note 39; Murray, *supra* note 16; Kevin Ashley, *Toward a Computational Theory of Arguing with Precedent*, Proc. of the Third Int'l Conf. on Artificial Intelligence & L. 93 (1991). However, the finding of similarity between cases is a prerequisite to any subsequent manipulations of the analogy.

113. V. Mital and L. Johnson, Advanced Information Systems for Lawyers 257 (1992). *See also* Celeste Tito, *Artificial Intelligence: Can Computers Understand Why Two Legal Cases Are Similar?*, 7 Computer/L.J. 409, 411-12 (1987). *But cf.* Burton, *supra* note 22, at 39 (stating that the process is nevertheless a mystery).

114. Mital & Johnson, *supra* note 113, at 257 (noting that the ruleless theory has been criticised because it would mean that people would not be able to say why or in what aspects two cases are similar).

115. Mital & Johnson, *supra* note 113.

116. *Id.*

(2) They must understand moral decision making.[117]

According to Tito, it is necessary to understand the analogue meaning of words to determine whether something is within a general category. Similarly, moral decisions must be made when determining at what level of generality things are to be compared.

While Tito says she is not interested in whether computers can mimic the results achieved by lawyers, but whether they can actually understand analogical reasoning,[118] her work does not consider the philosophical problems of what constitutes intelligence and understanding in computers.[119] Tito's work is still informative, however, if viewed as a discussion on the ability of computers to mimic the results achieved by lawyers.

A problem that faces all legal expert systems, including those that incorporate neural nets, is that they only model legal concepts. It is unavoidable that when an issue in the real world is to be considered by a computer, it has to be circumscribed by a limited number of factors. This circumscription will inevitably involve a loss of richness and the creation of a conceptual bias[120] in the computerized representation of the concept as compared to the real world concept. In Tito's conception, the computer only has a digital representation of concepts. Though this loss will be inversely proportional to, and dependent on, the complexity of the composition of the matrix used in the circumscription, if the input matrix does not accurately reflect the real world concept, then the conclusion drawn by the legal expert system will not be accurate.

It is as yet unclear whether the necessity of understanding moral decision making for the finding of similarity is a fundamental bar to computers performing analogical reasoning. Computers may yet be implemented that do this, though what this entails is presently unclear.

ii. Similarity and Neural Nets

In the quest to find similarity, neural nets can conceivably be used in several ways:

---

117. Tito, *supra* note 113. *See also* MACCORMICK, *supra* note 14 (arguing that the finding of an analog depends on the view of what purpose the legal system is trying to achieve). *Cf.* Sunstein, *supra* note 39, at 773-81 & n.116 (noting the need for a general theory with which to evaluate similarities and thinks this should cause skepticism about efforts to program computers to engage in analogical reasoning).

118. Tito, *supra* note 113.

119. For a concise discussion on this issue see the debate between John Searle, et al., *Artificial Intelligence a Debate*, SCI. AM. 262(1), 19 (1990) (providing a particularly interesting discussion of neural nets).

120. Karpf, *supra* note 107, at 299.

(a) By comparing matrices of factors;

(b) By determining weights to be given to factors that are used in other systems;

(c) By identifying new factors that are common to members of a group; and

(d) By determining similarity in a less reductionist fashion than the above.

For present purposes, approaches (a) and (b) are essentially the same. Both rely on matrices of factors being presented to a neural net. Although a neural net can classify patterns, deal with complex relationships and subtle variations in factors, and thereby determine similarity by determining how many attributes are shared, a key aspect of the finding of similarity has already been performed by the designer. The designer of the system has already made the all important decision as to what limited factors are to be considered relevant for a determination of similarity and, further, at what level of generality they are to be compared.

In this scenario, Tito's requirements mean that the computer can only find similarity at the level of attribute matching; more subtle aspects of similarity are outside the computer's scope. For this reason, systems such as PROLEXS that adopt the matrix approach will only ever have limited ability to reason analogically.

However, if a matrix can be chosen that can accurately model a real world concept,[121] then that matrix can be implemented in a neural net. This is a corollary of Kolmogorov's theorem.[122]

A key requirement in this approach is choosing the matrix used to represent the concept. But what factors are to be included? Neural nets could also conceivably be used to identify new factors that are common to a group. Bench-Capon shows how neural nets can find which factors are significant amongst noise,[123] but claims that the significance of these factors cannot be understood without independent knowledge of the domain.[124]

To say that the significance of such factors cannot be understood without prior domain knowledge, though, is not to say that the newly identified factors are not significant. According to some members of the critical legal studies movement, the reasons given in cases are not the whole list of reasons for the reaching of the results in those

121. "Accurately" is here being taken to mean: model to a degree of richness that is sufficient to satisfy lawyers.

122. *See supra* part II (discussing Kolmogorov's theorem).

123. Bench-Capon, *supra* note 56.

124. *Id.* at 296-97.

cases.[125] If this view of law is correct, then legal analysis and legal expert systems based solely on those decisions will not accurately reflect how and why cases are decided. Instead one should simply look at what actually occurs. Thus, when an analysis of a neural net highlights the importance of an unsuspected factor, this could be interpreted as telling us something important about the underlying legal domain. Consequently, the use made of noise is dependent on what jurisprudential theory the system's developers adopt; whether it is interpreted as a discovery about the law or is rejected as a technical anomaly.

The most promising approach to modelling similarities is the less reductionist approach taken by SCALIR.[126] Here, similarity is not judged solely on the presence or absence of specified factors, but also on the presence of subfactor information. Thus, even though two input matrices may share few factors at the conceptual level, they can still be regarded as similar if they directly or indirectly share common microfeatures. In this respect, SCALIR contains a closer approximation to employing the analogue meaning of words than do other systems.

However, before SCALIR type similarity determination can be implemented in a legal expert system, rather than solely a document retrieval system, the system's developers will have to choose how indirect a sharing of common microfeatures will amount to two objects being regarded as similar. This is equivalent to choosing at what level of generality the two objects are to be considered. Further, the approach adopted in SCALIR is still dependent on the system's designers choosing what concepts are to be used to model the legal domain. Thus, within Tito's framework it is still not possible to say the system implements moral decision making. However, in incorporating a closer approximation to the analogue meaning of words, the method to determine similarity adopted in SCALIR is more subtle than those in other neural net systems or that exist in symbolic reasoning systems.

It cannot be doubted, then, that neural nets can mimic the finding of similarity, though on a restricted basis. However, the accuracy of the similarity found will depend greatly on the composition of the matrix chosen by developers to describe the legal concepts.

iii. Open Texture

Two observations about the use of neural nets to resolve open texture can now be made. Since the similarity found by neural nets is

---

125. *See generally* Kennedy, *supra* note 25.

126. *See supra* part IV.B (discussing the operation of SCALIR).

crude compared to that achieved by humans, there is much scope for real world decisions to differ from those reached by neural nets because unconsidered factors will have been taken into account.[127] Secondly, legal analogical reasoning is not simply the finding of similarity between cases but involves manipulating the analogy found to achieve a desired result.[128] This is something that neural nets of themselves cannot perform. Consequently, by themselves neural nets have a limited ability to perform analogical reasoning. The ability of neural net systems to generalize input patterns and to perform a flexible form of similarity determination, however, makes them strong candidates for use in a hybrid analogical reasoning system

(b) Explanation and Justification

The use of neural nets as legal analogical reasoners faces a further problem. Mital and Johnson state that, "similarity cannot be thought of as an agent independent of the objects which are to be found similar; it may be said that it is more in the nature of a relation which the mind perceives after the fact."[129] Since similarity does not exist independently of our perception of it, creating this perception is of crucial importance. Unfortunately, this presents problems for neural nets. Presently, neural nets take a series of inputs and oracularly produce an output; it is left to the user of the system to infer why similarity was found.

Creating such a perception involves two things, explaining why the similarity was found and then justifying the finding. Several methods have been proposed to get explanations and justifications from neural nets, four of which are the following:

(1) Extract rules from the neural net;[130]

(2) Present to the user those nodes (factors) that had a positive contributory influence along with those that had a negative contributory influence on the decision;[131]

---

127. Walker, *supra* note 59, at 63; Birmingham, *supra* note 77, at 132-34.

128. *See generally* LEVI, *supra* note 22; MACCORMICK, *supra* note 14; BURTON, *supra* note 22; Gordley, *supra* note 39; Sunstein, *supra* note 39; Murray, *supra* note 16; Ashley, *supra* note 112.

129. MITAL & JOHNSON, *supra* note 113, at 256.

130. *Compare* Bochereau, *supra* note 80 (construing neural nets as being used specifically to extract rules) *with* David Warner, *A Neural Network-based Law Machine: the Problem of Legitimacy*, 2(2) LAW COMPUTERS & ARTIFICIAL INTELLIGENCE 135, 141 (1993) [hereinafter Warner, *The Problem of Legitimacy*] (arguing rules will be extracted at run time in response to questioning). It is possible that the latter approach would provide more flexibility.

131. Van Opdorp, *supra* note 57, at 285. *Cf.* Warner, *The Problem of Legitimacy*, *supra* note 130, at 139 (additionally attributing percentages of the input variables to the output variables).

(3) Present the training set of the neural net to the user;[132] and
(4) Create a hybrid system where the output of the neural net is explained ex post facto by other systems.[133]

The essential purpose of providing explanation and justification is to convince the human end user of the correctness of the result achieved and, in this respect, the intended audience and use of the system must be remembered.[134]

Stephen Gallant has given a detailed analysis of how rules can be extracted from neural nets,[135] though as Bench-Capon notes, we cannot be sure of the correctness of any rules derived from a neural net unless we have prior knowledge about the domain itself.[136]

However, while rules may provide an explanation of a result, it is hard to regard them as a justification. If a domain expert were asked, "How did you reach that conclusion?," a first answer might be, "It just came to me." Pressed further, the new response might be, "Factors X, Y and Z were present and this points to that result." A neural net can give a similar explanation by saying, "Factors X, Y and Z were present and this points to the result because they achieved that result in other cases." The expert (or neural net) might go further and formulate this last response with a rule such as, "Whenever factors X, Y and Z are present, then this result was achieved." As an explanation, this seems satisfactory: it was because of experience that the expert and neural net gave that result. A search for a more detailed explanation from a neural net, if even possible,[137] seems unnecessary.

Asking why the result is justified is different. What amounts to sufficient justification for a decision depends on the jurisprudential theory of law to which one subscribes. If one regards as justified a decision based solely on the fact that such a decision was reached in past situations, then "if . . . then . . ." rules as discussed above may be accepted as both explanation and justification; they are simply a shorthand way of saying this. However, if one's jurisprudential theory requires a more detailed justification, then it remains an open question whether a neural net can justify its results. Detailed justification may be possible using other systems, although a prerequisite is the adoption of a jurisprudential theory on what amounts to justification.

132. Van Opdorp, *supra* note 57, at 285.

133. *Id.*

134. Lambert and Grunewald, *supra* note 46; ZELEZNIKOW AND HUNTER, *supra* note 3, at 273-75.

135. GALLANT, *supra* note 44, at ch. 17.

136. Bench-Capon, *supra* note 56, at 296.

137. *E.g.*, DANIEL DENNET, CONSCIOUSNESS EXPLAINED 84-95 (1991) (arguing that these are aspects of human action that humans cannot themselves explain).

Proposals (b) and (c) for achieving explanations and justifications from neural nets are slightly different. In both cases, it is simply left to the user to infer why the information presented justifies the result achieved. The PROLEXS team stated that these approaches have not proved satisfactory.[138] Proposal (d), that of justifying the output of a neural net *ex post facto,* has not yet been reported as implemented though theoretical work is underway.[139]

Thus, it can be seen that neural nets have a limited ability to justify their results.

4. Inducing Rules

The work of Bochereau et al. and Bench-Capon have demonstrated that it is possible to extract rules from a trained neural net. Such rule induction, though, suffers the same problems as symbolic rule induction systems.[140] Essentially, all such rule induction is based on a statistical analysis of the underlying data. Thus, if the data is not statistically representative, then any rules induced could be spurious.

However, the flexible notion of similarity able to be embodied in a neural net may make neural nets more useful rule-induction systems than are current symbolic systems. As with analogical reasoning, a key step in inductive reasoning is the finding of similarity between cases so as to found a general rule. In traditional systems, such similarity is simply based on the presence or absence of factors chosen by the system designer. As discussed above, SCALIR embodies a more flexible notion of analogy than simple attribute matching. For this reason, a rule-induction system adopting SCALIR's concepts might be able to create relations and, thus, rules that would not be found with a symbolic system. It is conceivable that the more flexible approach to analogy embodied in SCALIR would improve rule induction. No work, however, has been undertaken on this point.

5. Compromise

Systems that model conflicts between rules by using compromise were discussed in part four. However, the use of neural nets to model contradictory rules is problematic. Philipps states that, "The neurons strive for equilibrium, and when the conditions of the equilibrium are translated into the terms of the case, the resulting solution cannot be

138. Van Opdorp, *supra* note 57, at 285.

139. *See* ZELEZNIKOW & HUNTER, *supra* note 3, at 273-75 (discussing the SPLIT-UP system).

140. *Id.*

totally unjust."[141] The equating of justice with compromise is questionable. Firstly, the two rules that were balanced may violate principles of formal justice, or they might offend against moral principles, in which case the resulting compromise cannot be said to be just. More fundamentally, justice does not necessarily equate with compromise. If justice is understood as meaning, "The result that a court of law would reach," then equating justice with compromise is unsupportable. Courts do not always achieve a result that is a compromise of the presented claims.[142] The point is not that compromise is never just or that what a court of law would do is just, only that in equating justice with compromise, a jurisprudential statement is being made that requires support. Thus, attempting to deal with conflicting rules through the use of compromise is not necessarily a desirable path. It depends on one's theory of justice.[143]

Used in the manner of Philipps, neural nets can only deal with contradiction through compromise.[144] Thagard's ECHO[145] has the potential to overcome this difficulty, as it does not model conflict through compromise. However, ECHO has problems of its own,[146] not least of which is the complexity of its representations. Since Thagard has not given detailed discussion of the legal use of ECHO, the possibility of using this system to model conflicting legal rules remains to be explored.

6. SCALIR Learning

Similar jurisprudential considerations arise from the proposal to make SCALIR learn from its users.[147] As an information retrieval tool this seems reasonable, any reasoning will be performed by the lawyer using the retrieved documents. However, if SCALIR were to be incorporated as part of a larger legal expert system, then such learning may not be justifiable.

141. Philipps, *supra* note 70, at 999.

142. *E.g.*, The Queen v. Watson; *Ex parte* Armstrong 136 C.L.R. 248, 249 (1976) (demonstrating that the High Court was forced to reject totally one line of authority when faced with conflicting lines of authority as to what amounted to judicial bias).

143. *E.g.*, MITAL & JOHNSON, *supra* note 113, at 259 (indicating that a conflict between the interpretation of factors within a case may be solved by the use of compromise).

144. *See infra* V.B.2 (discussing the practical implications of the necessity of ensuring the neural net is not presented with contradictory input).

145. *See supra* part IV.

146. ECHO requires competing hypothesis to be entered into the neural net along with their supporting facts. How each hypothesis is supported by evidence and what evidence contradicts what hypothesis then has to be entered by the system designer. Such decisions can be highly subjective and it is unclear what implications these requirements could have on the use of ECHO in the legal domain.

147. Rose & Belew, *A Connectionist and Symbolic Hybrid*, *supra* note 85, at 20-22.

Under Rose and Belew's proposal, learning in SCALIR would alter the very representation of the documents in the system. Consequently, a legal expert system adopting this approach would potentially alter its representation of the law each time it was used. Such a scenario has elements of the critical legal studies view that law is whatever we make it.[148] Aspects of this view may be true in the case of real lawyers and real judges. However, it is slightly bizarre to extend this to a legal expert system which has no direct affect on actual legal outcomes.

7. Normative Reasoning

Finally, it has been suggested that neural nets cannot model the legal decision making process because they cannot apply norms.[149] This is debatable.

If it is meant that neural nets cannot apply norms because of their normative content, this is incorrect. To the extent that norms can be expressed in terms of cases or rules, they can be modelled using a neural net. Any normative content in these cases or rules is irrelevant for this purpose. Indeed the very basis on which neural nets operate can be viewed as an application of the norm that like cases should be decided alike.

If it is meant that norms cannot be expressed in terms of cases, but must be represented as rules, then it still cannot be accepted that neural nets cannot model legal decisions. It is possible that localist neural nets[150] can be used to model norms.

If it is meant that neural nets cannot apply norms because they have no normative content for the neural net itself, then this is also debatable. This is tied to the question of whether neural nets and computers can think, which, though beyond the scope of the paper, is still an open question.[151]

However, it may be true that the result received from a neural net cannot force a value decision. This touches on moral philosophical questions that are also beyond the scope of this paper.

---

148. Kennedy, *supra* note 25.

149. MITAL & JOHNSON, *supra* note 113, at 253.

150. Neural nets can be classified not only according to their learning rules and architectures, but also as distributed or localist networks. In distributed networks, of which adaptive filter networks are one type, only the nodes at the input and output levels represent real world concepts, hidden layers are simply there to aid in the mapping performed by the network. In localist models, each node of the network represents a real world concept.

151. Searle, *supra* note 119.

B. *Methodological Concerns in the Use of Neural Nets*

In addition to the general jurisprudential issues associated with neural nets so far discussed, the manner in which neural nets are actually implemented in legal applications has implications for the jurisprudential theory embodied in the system.[152] The two most troubling aspects of the many uses discussed in the previous part are the use of hypotheticals to train neural nets and the manner with which contradictory data is dealt.

1. Hypotheticals

Neural nets rely on statistical analysis of the underlying data presented during training. Thus, to create accurate models, they require data that is statistically representative.

In their discussion of the uses of neural nets in law, Mital and Johnson note that much of the law remains unreported and that neither all possible nor all anticipated situations have been covered even by unreported cases.[153] This presents significant problems for the creation of neural nets in law. As outlined in the introduction to neural networks, neural nets create generalizations from the information presented to them during their training. The quality of these generalizations (in respect to the degree they reflect the actual outcomes of cases) is dependent on the cases used to train the network. A lack of cases will lead to spurious generalizations.[154]

The lack of reported cases with which to train neural nets has been reported by a number of researchers.[155] In an attempt to overcome this problem, researchers have resorted to creating hypothetical cases with which to train their neural nets.[156] It must be realized however, that once a neural net has been trained with any hypotheticals,

152. Apart from the discussion of the use of hypotheticals and compromise which follows, it should be noted that the study of neural nets is still a comparatively embryonic field. A multitude of network designs exist from which application developers can choose. Designs generally have numerous design parameters the values of which can be chosen more or less ad hoc. Both the type of network design used and the values of the parameters chosen, affect the behaviour of the neural net. How many hidden layers to include in the neural net and how many nodes to include in each of those layers, the learning rule and learning rate, the amount of noise present when the network is trained and even the order in which training examples are presented to the neural net can all affect the neural net's behaviour and the classifications it produces. Thus these factors can affect the way the neural net reasons with the information presented to it. However, the legal literature discussing neural nets does not discuss such issue, with the exception of Rose & Belew, *supra* note 85.

153. MITAL & JOHNSON, *supra* note 113, at 265.

154. Van Opdorp, *supra* note 57, at 282-84.

155. *E.g.*, Karpf, *supra* note 107; Hobson and Slee, *supra* note 54; Walker, *supra* note 59.

156. Hobson and Slee, *supra* note 54, at 12; Walker, *supra* note 59, at 57. *Cf.* Bench-Capon, *supra* note 56 (discussing use of hypotheticals to train the author's ANN within an entirely

then, unless one subscribes to a critical theory of law, it is no longer a system that reasons solely with the law.[157] It is an amalgam of the law and expert belief. This may or may not be problematic, depending on the purpose the system is designed to achieve.

Training with hypotheticals is said to incorporate the heuristic knowledge of the domain expert.[158] It seems justifiable to argue that predicted case outcomes generated by an expert do incorporate heuristic knowledge about how an expert would reason with a sparse set of cases, but the conclusion remains that any result reached by the system is not based solely on the law.

If it is decided to use hypotheticals, it seems necessary that they at least be generated by a domain expert. In Bench-Capon's neural net, all the hypotheticals were generated by another computer program.[159] At the very least, the rules in this program should be generated by a domain expert. But in such a case, the question arises why these rules are not simply incorporated into a symbolic reasoner.

Philipps has argued that his neural net need only be trained with ten training examples, as long as they are "prototypical."[160] However, while this may be true in the case of the simple rules that were there modelled, it seems difficult to apply to neural nets trained with cases and designed to resolve open texture.

The lack of training data needs to be addressed if neural nets are to be created that reason with the law.

2. Contradictory Input Data

A further problem facing the use of neural nets in law is the way in which they model conflicting data. Neural nets model conflicts in data by reaching a compromise between the conflicts. As previously argued, this may or may not be jurisprudentially acceptable.

If modelling conflict through compromise is jurisprudentially unacceptable, then in the training of the neural net it is extremely important to ensure that contradictory examples are not included in the training set. This is a huge difficulty. It is unclear how to determine whether two cases are conflicting if those cases differ in more than

---

hypothetical legal domain). *But see* Karpf, *supra* note 107, at 299 (criticizing the use of hypotheticals).

157. Similar observations can be made of case based reasoners. There is a large element of expert opinion as to what the relevant factors in a domain are, which cases are to be included in the knowledge base and whether those cases do or do not contain those relevant factors.

158. Van Opdorp, *supra* note 57, at 285.

159. Bench-Capon, *supra* note 56. *Cf.* Hobson and Slee, *supra* note 54 (making no mention of how hypotheticals are generated).

160. Philipps, *supra* note 70, at 995-996.

one aspect. Unless techniques can be created to determine where contradictions exist in a training set, doubt must be cast on the practical possibility of using neural nets to reason with the law.

Finally, in contrast to the work of researchers in legal expert systems and legal information retrieval systems, who are concerned with the practical uses of neural nets, several jurisprudes have also proposed using neural nets as a metaphor for the operation of law in society; as a representation of the interaction between the legislature, the courts and citizens.[161]

A detailed description of such a neural net theory of law has been given by Alexander Silverman,[162] who sees the law as a huge neural net in which:

> The judges and other legal actors are nodes of the network; the published case reports and statutes, teaching in the law schools, continuing education courses and learning on the job, and the informal and formal oral communications among the members of the legal community are the connections between nodes; the cases and statutes themselves are the patterns presented to and learned by the network.[163]

This is a descriptive theory of law[164] that sits between positivist and critical theories of law. The law is not the application of objective facts, but nor is it merely the preferences of individual judges.[165] Instead, no single actor or single rule determines the outcome of a case; the outcome emerges from the interaction of the whole system.[166] Similarly, under this theory rules and theories of law are to be regarded only as approximations of the underlying law, much as a neural net constructs a mathematical function to approximate the distinctions present in its input data.[167]

While the practical implementation of such a neural net is far beyond current capabilities, this is not Silverman's aim. According to Silverman, "At the most general level, our metaphor of law matters . . . new metaphors of law can lead to an increased awareness of alter-

---

161. Warner, *Role of Neural Networks*, *supra* note 69, at 138 (claiming that this view is inherent in the works of Anthony D'Amato).

162. ALEXANDER SILVERMAN, MIND, MACHINE, AND METAPHOR: AN ESSAY ON ARTIFICIAL INTELLIGENCE AND LEGAL REASONING (1993).

163. *Id.* at 80.

164. *Id.* at 81.

165. *Id.* at 80.

166. *Id.* at 80, 84-86 (explaining that an expanded version of this theory sees law not only as an interconnected network, but also as an interconnected network that 'resonates' with society; the law both influences and is influenced by the society in which it is constructed).

167. *Id.* at 81-83.

natives for the legal system."[168] With a new metaphor, the way we think about judges, law, society and our role therein, can radically change.

VI. Conclusion

Resurgent interest in neural nets has resulted in various applications in law. While neural nets do not have the potential to solve all problems present in current efforts to computerize legal knowledge, the use of neural nets does offer potentially great benefits in both the creation of legal expert systems and legal information retrieval systems.

Most promising is the ability of neural nets to aid in the determination of similarity between cases. The finding of similarity is a key step in the process of legal reasoning. Any legal expert system that seeks to model legal knowledge has to incorporate a means to determine similarity.

Neural nets offer a model of similarity that is more flexible than those found in existing symbolic reasoning systems and so have huge potential for use in legal expert systems.

Neural nets potentially offer other benefits, such as a means to model conflicts in rules and cases. Their ability to learn information adds further attraction to their use.

However, using neural nets to model conflict and to learn information has jurisprudential implications. The need for statistically significant numbers of cases with which to train neural nets, the jurisprudential implications of using hypotheticals during training, the need to ensure that training data is not contradictory and the currently limited ability of neural nets to justify their responses, all limit the present usefulness of neural nets in legal expert systems. While techniques have been proposed that potentially overcome both problems of contradiction and the ability to justify conclusions, little work has actually been conducted on these techniques. It must be ensured that the jurisprudential implications associated with these limitations do not undermine the overall project in which the use of neural nets is playing a part.

While neural nets can offer a new metaphor for law, it is only through future research that creates hybrid neural net/symbolic reasoning systems that we will truly be able to use computers to test the implications of our current jurisprudential theories.

168. *Id.* at 94-95.