1991

What is Legal Therapy and What is It Doing in the Curriculum - A View from Contracts

Robert Birmingham

University of Connecticut School of Law

WHAT IS LEGAL THEORY AND WHAT IS IT DOING IN THE CURRICULUM? A VIEW FROM CONTRACTS

by Robert Birmingham*

I am not sure that I know what you mean when you ask me if I don't believe that the law is a monad.
    —Holmes to Wu[1]

'Everyone maximizes utility.' This claim needs interpretation, because it looks filled with content, but really is innocuous, a tautology, nothing to get upset about. It means merely that one's behavior invariably maximizes *some* function. As Samuelson's imperative, "*[C]herchez la* maximization,"[2] suggests, often it is an interesting task and something of an art to recover the function from the behavior. I offer three examples of maximizing, eliminating purpose by degrees.

A places an amount of money in an envelope and seals the envelope. Then she tosses a fair coin. If the coin lands heads, she puts *twice* that amount of money in a second envelope and seals it. If the coin lands tails, she puts *half* that amount of money in the second envelope and seals it. She hands an envelope each to *B* and *C*. Before *B* opens her envelope, should *B* prefer to exchange envelopes with *C*? What if she opens her envelope and finds $10?

FIGURE 1
Test Question

(1) Landes and Posner recently endorsed "the hypothesis that the common law of torts is best explained as if the judges who created the law through decisions operating as precedents in later cases were trying

---

* Professor of Law, University of Connecticut; A.B., J.D., Ph.D., University of Pittsburgh; LL.M., Harvard University.
    1. Letter from O.W. Holmes to John C.H. Wu, Feb. 5, 1923, *reprinted in* JUSTICE HOLMES TO DOCTOR WU: AN INTIMATE CORRESPONDENCE, 1921-1932, at 7, 7 (n.d.). The correspondence at least on Holmes' side which alone is reported is not very intimate and consists of Holmes politely saying he does not know what Wu is talking about. *Cf.* CASES AND MATERIALS ON JURISPRUDENCE (J. Wu ed. 1958).
    2. Weintraub, *Book Review*, 28 J. ECON. LIT. 1192, 1192 (1990), *quoting* JOHN VON NEUMANN AND MODERN ECONOMICS 112 (M. Dore, S. Chakravarty & R. Goodwin eds. 1989).

to promote efficient resource allocation."[3] I.e. what delighted the old judges was approaching Pareto optima, although they could not recognize, as Landes and Posner see retrospectively, the source of their utility. (2) Polar bears are white. What are they maximizing? Bishop Montefiore says 'Nothing': "As for camouflage, this is not always easily explicable on neo-Darwinian premises. If polar bears are dominant in the Arctic, then there would seem to have been no need for them to evolve a white-colored form of camouflage."[4] Dawkins paraphrases,

> I personally, off the top of my head sitting in my study, never having visited the Arctic, never having seen a polar bear in the wild, and having been educated in classical literature and theology, have not so far managed to think of a reason why polar bears might benefit from being white,[5]

then says polar bears maximize sneaking up on seals. (3) A rope suspended between two points "will hang in the shape of a catenary, $y(x) = a_1 e^{\lambda x} + a_2 e^{\lambda x}$, because even a dumb rope knows that such a shape will minimize its center of gravity," and attaining a low center of gravity is just what a rope wants.[6]

I

If the extension of 'everyone' includes ropes, it includes law teachers: they too must implicitly maximize some function, the challenge being to recover it. To do so, I presuppose they want to teach law effectively, and ask, 'What understanding of legal theory makes their behavior maximizing?' The place of legal theory in the curriculum, the topic of this panel, falls out of the pragmatics. The questions of Figure 1 invites the conferee in the course of solving them to identify the theories underlying her behavior.[7]

Of course law teachers do not act alike. I read 'law teachers' restrictively, indeed parochially, making its extension not just educators in American law schools—the only ones I can introspect about—but mainly C. C. Langdell. The restriction has its disadvantages, since,

---

3. W. LANDES & R. POSNER, THE ECONOMIC STRUCTURE OF TORT LAW 1 (1987).

4. R. DAWKINS, THE BLIND WATCHMAKER 38 (1986), quoting H. MONTEFIORE, THE PROBABILITY OF GOD (1985).

5. *Id.*

6. Samuelson, *A Catenary Turnpike Theorem Involving Consumption and the Golden Rule,* in 3 P. SAMUELSON, THE COLLECTED SCIENTIFIC PAPERS OF PAUL A. SAMUELSON 93, 93 (1972).

7. Nalebuff, *Puzzles: The Other Person's Envelope Is Always Greener,* 2 J. ECON. PERSPECTIVES 149, 150 (1988), 3 J. ECON. PERSPECTIVES 171 (1989).

starting with Holmes, who called him "perhaps the greatest living legal theologian,"[8] almost everyone, e.g. Gilmore, although not Grey,[9] has laughed at Langdell, Gilmore saying, "[I]f Langdell had not existed, we would have had to invent him. Langdell seems to have been an essentially stupid man who, early in his life, hit on one great idea to which, thereafter, he clung with all the tenacity of genius."[10]

Okay, but *even Gilmore* believed Langdell matters, continuing: "However absurd, however mischievous, however deeply rooted in error it may have been, Langdell's idea shaped our legal thinking for fifty years."[11]

Still, Gilmore is careless about what idea Langdell clung to: that law is a science, surely; yet Gilmore has no understanding of what a science is. A discipline whose propositions are conceivably refutable (Popper)? No matter: our concern is the pedagogy accompanying Langdell's idea, described thus:

> Langdell's chief innovation in legal education was the introduction of the so-called case method of teaching in which the principal (in Langdell's original version, the only) materials presented to the student are the reports of decided cases, whose meaning is to be worked out by study and in classroom discussion. The case method, which was bitterly attacked for a generation after Langdell had introduced it, had by the time of World War I, been adopted in almost all American law schools. Langdell's *Cases on Contracts* was the first casebook of all.[12]

To get anywhere, we must inspect the case method in action. Nevertheless as far as I know there is no transcription of Langdell's class. We are told, "Teaching his first class in Contracts" in 1870, Langdell, newly Dean of the Harvard Law School, "began not with the customary introductory lecture, but by asking 'Mr. Fox, will you state the facts in the case of *Payne v. Cave*?' "[13] We are not however told what Fox replied etc. Nor does a law teacher today commence with *Payne v.*

---

8. Book Notice, 14 AM. L. REV. 233, 234 (1880) (reviewing C. LANGDELL, A SELECTION OF CASES ON THE LAW OF CONTRACTS, WITH A SUMMARY OF THE TOPICS COVERED BY THE CASES (2d ed. 1879)).

9. Grey, *Langdell's Orthodoxy*, 45 U. PITT. L. REV. 1 (1983).

10. G. GILMORE, THE AGES OF AMERICAN LAW 42 (1977).

11. *Id.*

12. *Id.* at 125 n.3.

13. Grey, *supra* note 9, at 1.

Cave[14], a case no longer read, so that even Gilmore's casebook,[15] the most preserving of the traditional cases, omits it. No unbearable loss: in *Payne*, Cave at auction bid for a pewter worm but withdrew his bid before the hammer fell; the court held he might do so. The case is fun to teach only because of the stated great weight of the worm, "1300 hundredweight," and because few students know what a worm *is*: certainly not the "small harmless serpent that lives in the earth" or the "grub[] that gnaw[s] wood and furniture" as Dr. Johnson defines it.[16] It is a spiral condensing tube for distilling;[17] still, the weight is unexplained.

Hence I turn to fiction, in the first instance Osborne's *The Paper Chase*.[18] The novel begins with Kingsfield doing push-ups in his office, but presently he goes to Contracts class, at which point we pick up the action.

> Most of the first year students, in anticipation of their first class at the Harvard Law School, were already seated as Professor Kingsfield, at exactly five minutes past nine, walked purposely through the little door behind the lecture platform. He put his books and notes down on the wooden lectern and pulled out the seating chart. One hundred and fifty names and numbers . . . .
>
>   . . . .
>
> At exactly ten past nine, Professor Kingsfield picked a name from the seating chart. The name came from the left side of the classroom. Professor Kingsfield looked off to the right, his eyes following one of the curving benches to where it ended by the window.
>
> Without turning, he said crisply, "Mr. Hart, will you recite the facts of *Hawkins* versus *McGee*?"
>
> When Hart, seat 259, heard his name, he froze. Caught unprepared, he simply stopped functioning. Then he felt his heart beat faster than he could ever remember its beating and his palms and arms break out in sweat.
>
> Professor Kingsfield rotated slowly until he was staring

---

14. 3 Term. R. 148, 100 Eng. Rep. 502 (K.B. 1789).

15. F. KESSLER, G. GILMORE & A. KRONMAN, CONTRACTS: CASES AND MATERIALS (3d ed. 1986).

16. S. JOHNSON, A DICTIONARY OF THE ENGLISH LANGUAGE N.P. (London 1755).

17. WEBSTER'S NEW INTERNATIONAL DICTIONARY 2953 (2d ed. 1960).

18. J. OSBORN, THE PAPER CHASE 73 (1971).

down at Hart. The rest of the class followed Kingsfield's eyes.

"I have got your name right?" Kingsfield asked. "You are Mr. Hart?" He spoke evenly, filling every inch of the hall.

. . . .

[Hart's] voice floated across the classroom: "I . . . haven't read the case. I only found out about it just now."

Kingsfield walked to the edge of the platform.

"Mr. Hart, I will myself give you the facts of the case. *Hawkins* versus *McGee* is a case in contract law, the subject of our study. A boy burned his hand by touching an electric wire. A doctor who wanted to experiment in skin grafting asked to operate on the hand, guaranteeing that he would restore the hand 'one hundred percent.' Unfortunately, the operation failed to produce a healthy hand. Instead it produced a hairy hand. A hand not only burned, but covered with dense matted hair.

"Now, Mr. Hart, what sort of damages do you think the doctor should pay?"

. . . .

"The difference between what he was promised and what he got, a worse hand?" Hart asked.

Kingsfield stared off to the right, picked a name from the seating chart.

"Mr. Pruit, perhaps you can tell the class if we should give the boy the difference between what he was promised and what he got, as Mr. Hart suggests, or the difference between what he got, and what he had."[19]

Kingsfield, like Fuller[20] before him, starts with *Hawkins v. Mc-Gee.*[21] Hamilton, Rau and Weintraub still do.[22] The exchange is hard on Hart. Regard, however, how Kingsfield handles Hart's answer. The problem for Kingsfield is that improbably Hart has already hit on the expectation measure of damages: Hart has given Kingsfield the correct answer. The exigencies of plot may require this. Hence Kingsfield moves on to Pruit, hoping he will answer incorrectly, getting the class barking down the trail of reliance damages. It is a good thing pedagog-

---

19. *Id.* at 6-9.

20. L. FULLER, BASIC CONTRACT LAW 1 (1946).

21. 84 N.H. 114, 146 A. 641 (1929). The 'before' is logical not chronological, Kingsfield having started teaching before Fuller's casebook.

22. R. HAMILTON, A. RAU & R. WEINTRAUB, CONTRACTS: CASES AND MATERIALS 1 (1984).

ically that Pruit be wrong.

I take my second exhibit from Turow's *One L*.[23] We are still at Harvard, but the Contracts class is further along. Perini not Kingsfield is teaching.

> We were studying *Hadley v. Baxendale*, a famous case which established a limit on the kinds of damages a winning plaintiff in a contract suit could collect. Perini asked us what the rule of Hadley [*sic*] was *not* designed to do. He said there was a one-word answer. People raised their hands offering responses ranging from "work" to "make sense," and Perini toured the room, quickly shooting them down: "No," "Never," "Silly," "You think *that* makes sense?"
>
> When he saw my hand, he whirled and pointed.
>
> "To punish," I said. I was shocked I was speaking. My heart was slamming in my chest.
>
> Perini came closer, tilting his head. "How so?"
>
> "The way the rule works, it doesn't act to punish somebody who breaches a contract."
>
> "What difference does that make?"
>
> "It means that damages aren't awarded to deter breach."
>
> "What are they intended to do, then?" Perini asked.
>
> "Just compensate the loss."
>
> "Right!" said Perini. "Contract damages are merely intended to compensate plaintiff for his loss. You leave all that soul-splitting over punishment behind in Torts and Criminal Law—it's not for Contracts!"[24]

MAGEE But [logical positivism' must have had real defects. What do you now, in retrospect, think the main ones were?

AYER [the leading Anglo-American logical positivist] Well, I suppose the most important of the defects was that nearly all of it was false.

FIGURE 2
All About Logical Positivism

Perini screws up, reminding the reader of Wittgenstein's remark: "It is his enormous mistake which is great. I.e. roughly speaking if you just add a ' ~ ' to the whole book it says an important truth."[25] Perini is almost that badly off, be-

---

23. S. TUROW. ONE L (1977).

24. *Id.* at 73.

25. L. WITTGENSTEIN. LETTERS TO RUSSELL. KEYNES. AND MOORE 159 (G. von Wright ed. 1974).

cause a better answer than 'Punish' is 'Compensate'. Hence he is teaching them all wrong substantively. The method is what matters here, however: the dialogue occurs respecting a case, *Hadley*,[26] that every common lawyer knows, about a broken mill shaft returned to the manufacturer as a model for a replacement. The carrier in breach of contract sent the shaft by boat instead of train, delaying its arrival. The mill had been stopped for want of a shaft and Hadley sought lost profits. He lost because (despite the headnote) he had not told the carrier or at least Baxendale himself that the mill was stopped. No expectation damages here, not because they are not efficient, but because a party to a contract contemplating breach can only tell whether to breach if she knows the other party's prospective loss, and it is worth motivating this other party to tell her this loss at the cost of otherwise underdeterring breach. But, thinking pedagogically, notice how Perini likes getting wrong answers—although he fails to exploit the students' mistakes, merely replying 'Silly' and so forth.

That is the data. I am not all that concerned about the 'legal' part of 'legal theory'; 'theory' is the problem I address. The logical positivists blessed us with a definition of 'theory' everyone once accepted: "a set of sentences or propositions, expressible in the first-order predicate calculus."[27] That made a theory a linguistic entity. Ignore the 'first-order' part; it just requires that variables range over individuals, not sets. In a first order language, one can talk about a contract, the thing, but not the property of being a contract (contracthood). But lawyers do not need to do that. So cast about in the law for something that is a theory in this logical positivist sense. A good candidate is Restatement (Second) of Contracts,[28] the writing. Law students quickly catch on that the case method teaches legal rules only inefficiently, and buy *Gilbert*'s. Neither Kingsfield nor Perini direct, e.g., 'Learn sections 1-90'. Logical positivism however is defunct as Figure 2 indicates. The favored definition of 'theory' today likely is that a theory is not the sentences themselves but the underlying semantics.[29] Then the law would be not the Restatement but a domain, a function from 'contract' etc. to sets of sets of elements in the domain, ... The structure inter-

---

26. Hadley v. Baxendale, 9 Ex. 341, 156 Eng. Rep. 145 (1854).

27. Churchland, *On the Nature of Theories: A Neurocomputational Perspective,* in SCIEN-TIFIC THEORIES 59, 59 (W. Savage ed. 1990).

28. RESTATEMENT (SECOND) OF CONTRACTS § 265 (1981).

29. *See, e.g.,* R. GEIRE. EXPLAINING SCIENCE: A COGNITIVE APPROACH (1988); F. SUPPE. THE SEMANTIC CONCEPTION OF THEORIES AND SCIENTIFIC REALISM (1989); B. VAN FRAASSEN. LAWS AND SYMMETRY (1989). *Cf.* B. MAGEE. MEN OF IDEAS 131 (1978).

prets the Restatement. But Kingsfield and Perini are not directly teaching that either. Three other definitions of 'theory', more strictly legal, I dismiss by note.[30]

II

I look for a discipline with these characteristics. *A.* Its subject matter is intractable like law. *B.* The hardware (wetware) for processing it is similar. *C.* The teaching technique is the same. Interpreting sonar signals suffices.

A. *Intractability*

Clancy's novel *The Hunt for Red October*[31] describes these two protagonists aboard the USS *Dallas*, a 688-class attack submarine:[32]

Sonarman Second Class Ronald Jones

   Three years earlier, Jones had been asked to leave the California institute of Technology in the middle of his junior year. He had pulled one of the ingenious pranks for which Cal Tech students were justly famous, only it hadn't worked. Now he was serving his time in the navy to finance his return. It was his announced intention to get a doctorate in cybernetics and signal processing. . . . Jones' IQ was 158, the highest on the boat by a fair margin. He had a placid face and sad brown eyes that women found irresistible. On the beach Jones had enough action to wear down a squad of marines.

BC-10

   During her last overhaul, the *Dallas* had received a very special toy to go along with her BQQ-5 sonar system. . . . It was the most powerful computer yet installed aboard a submarine. Though only about the size of a business desk, it cost over five million dollars and ran at eighty million operations per second. . . . [T]he BC-10 stripped away ambient noise and other naturally produced sea sounds to classify and identify man-made noise. It could identify ships by name from their individual acoustical signatures, much as one could identify the finger or voice prints of a human.

An index of intractability is the reputed intelligence of the participants. In *The Hunt for Red October* peculiarity accompanies genius. So the

---

   30. (1) *The mean content of courses called 'Jurisprudence' or 'Legal Philosophy'.* The distribution here is likely bimodal. There are historical surveys, sort of stamp collecting, as in Christie's casebook. If jurisprudence is not largely taxonomic, usually it is Oxford ordinary language philosophy, as by H.L.A. Hart. Law schools train lawyers, who need not know these things. Legal philosophy done as ordinary language philosophy is speaking in a dead tongue. *Cf.* W. MILLER. A CANTICLE FOR LEIBOWITZ (1959).

   (2) *Law and . . . .* Entries here include Feminist Legal Theory, Law and Economics, Critical Legal Studies (read maybe insensitively as Law and Marxism), Law and Literature, Critical Race Studies. These are more nearly nonlegal theory applied to law. They—especially Law and Economics—may, however, *be* a way of doing law, as being a tree was once a way of being a person.

   (3) *Legal metatheory.* Calling the law a 'monad' is not straight theory but metatheoretical—talk *about* legal theory. Cases get decided identically whether or not law is a monad. This Essay is metatheoretical. The required demarcation, however, is between nontheory and theory, not between theory and metatheory.

   31. T. CLANCY, THE HUNT FOR RED OCTOBER (1984).

   32. *Id.* at 70-71, 74.

reader encounters Jones "in his usual trance."[33] And sonarpersons commune with whales (whose songs can travel through the water around the globe). BC-10 is smart too, although like a billion dollars, eighty MIPs is not what it used to be, and Jones one time calls BC-10 "a real *kludge*" for misidentifying a signal as a magma displacement (Clancy explains: "Jones' epithet was the most pejorative curse of electronics people"[34]). Nor does Clancy say how BC-10 does on a beach. Off hand, however, Jones and BC-10 seem smarter than Hart; and the narrator of *One L* did not make law review.
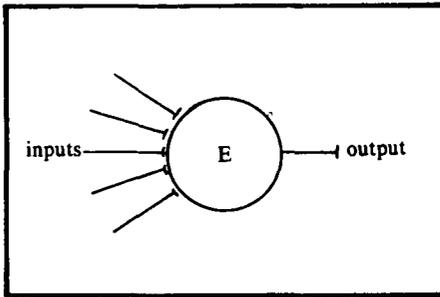
B. *Hardware*[35]

FIGURE 3
Processing Unit

The computer I describe, call it 'BC-11', is a massively parallel network. It is an artificial brain. Its function on a submarine is for my purposes to distinguish a rock from a mine, useful because mines but not rocks explode. I first specify the activities of its individual processing units (neurons), then indicate how they are connected into a network.

1. Processing Unit. Figure 3 portrays a processing unit or neuron of BC-11. Call a representative output of the jth processing unit 'p_j', and the weight of p_j as an input of the ith processing unit 'w_{ij}'. The bias of the ith unit, a constant added to the weighted sum of its inputs, is θ_i. Translate 'Σ' as 'sum of'. The ith processing unit, then, initially calculates E_i by Equation (1):

$$E_i = \sum_j w_{ig} p_j + \theta_i \tag{1}$$

For example, for two inputs, .2 and .7, with weights -.5 and 1, and a bias of .4, $E = .2(-.5) + .7(1) + .4 = .1$.

---

33. *Id.* at 70.
34. *Id.* at 76-77.
35. The mathematics is from Gorman and Sejnowski, *Learned Classification of Sonar Targets Using a Massively Parallel Network*, 36 IEEE TRANSACTIONS: ACOUSTICS, SPEECH, AND SIGNAL PROCESSING 1135 (1988). I do not read *IEEE Transactions: Acoustics, Speech, and Signal Processing* regularly. I got onto Gorman and Sejnowski through Churchland, *supra* note 27.

Then, to get its output, p_i, the ith processing unit transforms E_i by Equation (2):

$$p_i = P(E_i) = \frac{1}{1 + e^{-E_i}} \qquad (2)$$
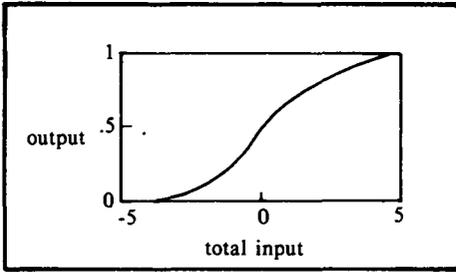
Here e is the base of the natural logarithms $\sim$ 2.7. If E_i = 0, then e^*-Ei = 1 and consequently p_i = $1/(1 + 1)$ = .5. As E_i increases positively, p_i gets larger, at first rapidly, then slowly, approaching 1; as E_i increases negatively, p_i correspondingly becomes smaller, approaching 0. Always then p_i lies between 0 and 1. The function P is a sigmoidal transformation; Figure 4 graphs it.

FIGURE 4
Sigmoidal Transformation

2. *Network.* BC-11 is a three-layer network of the individual processing units, connected as in Figure 5.[36] The input units receive data from outside the network, here registering the energy at various frequencies of sonar signals from rocks and mines, normalized to real numbers between 0 and 1. Each input unit is connected to each hidden unit, although I have not drawn all the connec-

FIGURE 5
BC-11

tions. The two output units together transmit outside the computer an ordered pair of numbers between 0 and 1. If the object listened to is a mine, ideally they signal <1, 0>; if a rock, <0, 1>. With 105 ($\sim 10^2$) connections in Figure 5, BC-11 is simpler than Jones, whose brain contains 10^{14} connections.
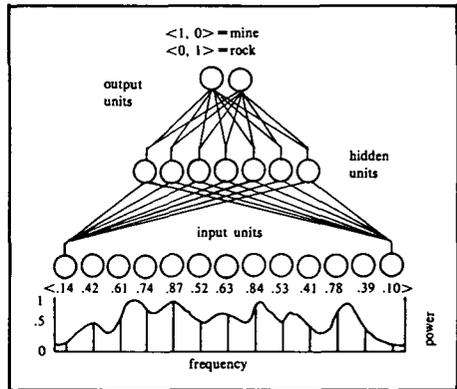
C. *Teaching BC-11*

One can teach the computer many skills besides differentiating be-

---

36. Churchland, *supra* note 27, at 75 fig. 5.

tween rocks and mines: e.g. to recognize phonemes (ā), hard because the sound patterns differ greatly among speakers; and to read out loud from a printed text. I stick with rocks and mines. Here the "difficulty is twofold: echoes from both objects sound indistinguishable to the casual ear, and echoes from each type show wide variation in sonic character."[37]

The teacher, call her 'Kingsfield' ('Kingsfield₂'), collects sample sonar signals from rocks and mines. They become the teaching set, a casebook's analogue. So Kingsfield gives BC-11 a signal (states a case). At the beginning, BC-11, like Hart, just guesses. In technical terms, their w_{ij}s have arbitrary small values. If BC-11 answers correctly, Kingsfield proceeds to the next signal. (In the legal event, Kingsfield sent the same signal to a different network: Pruit.) But outside fiction the computer will err as did Perini's students. Imagine Kingsfield offers a rock noise and BC-11 replies: '$<.7, .2>$'. Then Kingsfield sends an error message, the equivalent of Perini's "No," "Never," "Silly," or "You think *that* makes sense?"

Okay. So BC-11 knows it has made a mistake. It has an algorithm, a sort or recipe, that tells it what to do next. The word is from the arabic *al-Khuwārizmi*, surname of abu-Ja'far Mohammed ibn-Mūsa, the ninth-century author of a book on arithmetic, translated into Latin with the name *algorismus*.[38]

First, it locates the error, or rather distributes it among the outputs of its processing units. It does this in two steps. It calculates the error terms for its two output units, $\delta_i^{(N)}$, by *Equation (3)*,

$$\delta_i^{(N)} = (p_i{}^* - p_i) \ P'(E_i^{(N)}),\tag{3}$$

where

N = number of layers (three)
$p_i{}^*$ = desired output of ith unit
p_i = actual output of ith unit
$p_i{}^* - p_i$ = how far off is output of ith unit
$P'(E_i^{(N)})$ = slope of the function of Figure 4

Given that the signal is a rock's, $p_1{}^* - p_1 = 0 - .7 = -.7$ and $p_2{}^* - p_2 = 1 - .2 = .8$.

Then it backpropagates the error to the units of the other layers, that is, it figures out these units' complicity in the final errors, by

---

37. Churchland, *supra* 27, at 70.
38. WEBSTER'S NEW INTERNATIONAL DICTIONARY, *supra* note 17, at 64.

Equation (4),

$$\delta_i^{(n)} = \sum_j \delta_j^{(n+1)} \; w_{ji}^{(n)} \, P'(E_i^{(n)}) \tag{4}$$

Equation (4) directs: to obtain the error of the ith hidden unit, for $n = 2$, or of the ith input unit, for $n = 1$, weight the errors of the units it transmits to, those in the next layer up, by the importance of its signal to those units; add up the weighted errors; and, again, multiply by the slope of the function P.

Now BC-11 has calculated the errors by its various processors. Second, its the algorithm tells it to change the w_{ij}s according to Equation (5),

$$\Delta w_{ij}^{(n)} + \epsilon \, \delta_i^{(n+1)} \quad p_j^{(n)}, \tag{5}$$

the 'Δ' indicating the change, the 'ϵ' setting BC-11's rate of learning. If ϵ is too big, if the learning is insufficiently gradual, BC-11 might oscillate among misunderstandings. The process completed, Kingsfield goes on to another sonar signal (in Fuller's casebook, *Clark v. Marsiglia*,[39] whose plaintiff failed to mitigate damages, since he kept on cleaning paintings after their owner, breaking the contract, told him to cease). BC-11 will make more mistakes, receive error messages, and correct the w_{ij}s. So will Pruit. That is all there is to it. Notice however that a case contains its own error message, because an opinion not only states the facts but tells their legal consequences.

III

BC-11, and Hart, should eventually achieve w_{ij}s that elicit few error messages. The way to tell is to test them. Hence the nautical Kingsfield will give the machine a test set of sonar signals from rocks and mines, signals it has not heard before, but should have learned to interpret. For instance, the test question of Figure 1 might be answerable from experience

> Suppose you're on a game show and given a choice of three doors. Behind one is a car; behind the others are goats. You pick Door No. 1, and the host, who knows what's behind them, opens No. 3, which has a goat. He then asks if you want to switch to No. 2. Should you switch?

FIGURE 6
Teaching Question

with the teaching of Figure 6,[40] and itself belong to a later teaching set.

39. 1 Denio 317 (N.Y. 1845).
40. Letter from Robert Sachs, Ph.D., to Marilyn vos Savant, n.d., in vos Savant, *Ask Marilyn*, PARADE MAG., Dec. 2, 1990, at 25. *See* Tierney, *Behind Mority Hall's Doors: Puzzle*,

The test set for Hart, the Contracts examination, went thus, as far as we know, *The Paper Chase* being reticent about it, as if the topic were copulation, and the book respectably Victorian.

> Hart stood outside Ames Hall. The exam would begin in ten minutes. He kept away from the small groups of students nervously talking on the lawn. He kept a tree between himself and the others.
>
> He was talking to himself, not openly, but in his mind, addressing himself in commands.
>
> "Listen, Hart," he said, "shake contracts. Pour it out on that exam. Shake it loose and pour it all out. Don't leave the smallest fact in your mind. Make a clean sweep. Write it all down and sweep it out."
>
> At five of nine, he walked into the building, up the steps, taking them one at a time in measured strides, looking straight ahead. Into the building, down the hall, on the last walk.
>
> The exam books were laid out on the desks, an empty seat between each book so there would be no cheating. Hart took a seat at the side of the room so that there would be as few people around him as possible.
>
> Others were filing in too. Some were studying their outlines in a futile attempt to cram in the things they should have learned before. Others were just moving blindly toward the nearest seat in a short mindless dance, their arms hanging limp at their sides.

% Correct Classification		
Number of Hidden Units	Average Performance on Training Sets	Average Performance on Test Sets
0	79.3	73.1
2	96.2	85.7
3	98.1	87.6
6	99.4˙	89.3
12	99.8	90.4
24	100.0	89.2

FIGURE 7
Effect of Hidden Units (1)

> "O.K., baby, O.K.," Hart said, rubbing his palms together. The people near him yanked around.
>
> "O.K.," he repeated, "bring that fucking test in here."[41]

The next sentence is: "Hart walked out into the yard after his contracts

---

Debate, and Answer?, N.Y. Times, July 21, 1991, §1, at 1, col. 5.

41.   J. OSBORN, *supra* note 18, at 178-79.

exam."[42]

. BC-11 does learn, as Gorman and Sejnowski show for a network with sixty input units. The excitement for theory is in its hidden units. Figure 7 shows BC-11's response to training sets, then to test sets of unfamiliar signals of rocks and mines, as a function of the number of these units.[43] Figure 8 displays BC-11's learning curves, also as functions of these units.[44]

Studying the data tells one: (1) hidden units matter, BC-11's success rate on test sets varying between 73.1% and 90.4% as a function of the number of these units; (2) a computer learns faster if it has fewer hidden units, as one sees from Figure 8, where e.g. the left part of the curve for no hidden units is above the other curves; (3) if a computer has too many hidden units, it perhaps does perfectly on the training set, but its performance on the test sets falls off. Basically if it has too few hidden units it is too stupid to learn much; if it has too many, it memorizes the training set, hence cannot deal with new signals.
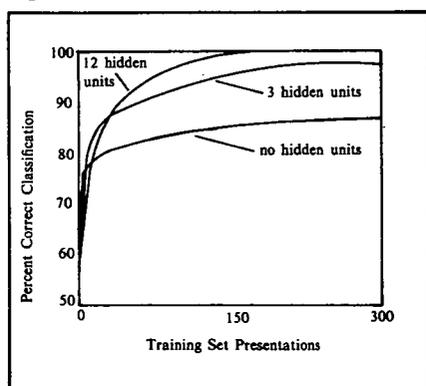

FIGURE 8
Effect of Hidden Units (2)

We have all encountered students using too few or too many units. (2) above, that students with suboptimally many units learn more quickly, is especially true in Contracts. It is a year course. Usually I must examine the students in December, because other students switch into my class for the second semester, and they are examined by the teachers they are leaving. I speculate that on a December examination in Contracts, I get a better reading by inverting the grades. Understanding strikes the Contracts student using the right number of neurons like light hitting Saul on the road to Damascus: suddenly, in April, she sees. BC-11 learns more gradually.

Keeping in mind from Part I that the usual definitions of 'theory' did not fit Kingsfield's and Perini's practice, think of BC-11 having a

---

42. *Id.* at 179.

43. Gorman and Sejnowski, *supra* note 35, at 1138 table 1.

44. *Id.* at 1139 fig. 4. The data differ somewhat between the figures because the training sets of Figure 7 are selected to include all target angles, while those of Figure 8 are randomly selected.

theory that *is* the w_{ij}s.

> [N]etworks have contrived a system of internal represen-
> tations that truly corresponds to important distinctions and
> structures in the outside world, structures that are not explic-
> itly represented in the corpus of their sensory inputs. The
> value of those representations is that they and only they allow
> the networks to 'make sense' of their variegated and often
> noisy input corpus, in the sense that they and only they allow
> the network to respond to
> those inputs in a fashion
> that systematically
> reduces the error
> messages to a trickle.
> These, I need hardly re-
> mind, are the functions
> typically ascribed to
> *theories.*
>
> . . . An individual's
> overall theory-of-the-
> world, we might venture,
> is not a large collection or
> a long list of stored sym-
> bolic items. Rather, it is a specific point in that individual's
> synaptic weight space. It is a configuration of connection
> weights, a configuration that partitions the system's activa-
> tion-vector space(s) into useful divisions and subdivisions rela-
> tive to the inputs typically fed the system. 'Useful' here means
> 'tends to minimize the error messages'.[45]

FIGURE 9
BC-11's Theory

The outputs of the hidden units are a vector in a space having as
many dimensions as there are hidden units. Figure 9 shows this space
for three dimensions. The w_{ij}s partition that space into a part for which
BC-11 responds 'rock' and a part for which it responds 'mine'. Within
these parts are regions of paradigmatic rock and mine vectors respec-
tively, represented by the hatched areas of Figure 9. BC-11's implicit
theory is: only mines are metallic.[46]

With 'theory' so defined, there is *invariably* a theory. The role of
legal theory in the curriculum is pervasive. "[N]o cognitive activity
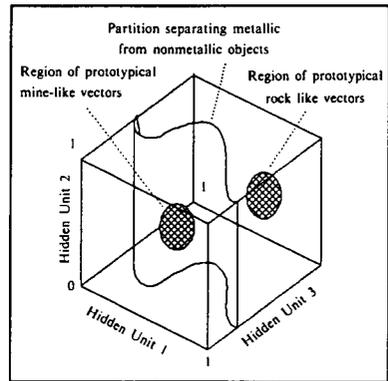
---

45. Churchland, *supra* note 27, at 82.
46. *Id.* at 80.

whatever takes place in the absence of vectors being processed by some specific configuration of weights"; hence "no cognitive activity takes place in the absence of some theory or other."[47] And if the cognitive activity is legal, so is the theory legal theory.

IV

The logical positivists, besides thinking of theories as sentential, also distinguished between observational terms and theoretical terms. In the end, 'liability' belongs to the observation language; but 'contract' is incorrigibly theoretical.

The legal version of the conundrum goes back at least to Holmes[48] who insisted the bad person wanted from her attorney only advice about liability: what the law would do to her as a function of her behavior. Also it has been common philosophically to deny contracts exist, because they are not concrete like cows: a contract lacks causal powers, in contrast to a cow, which can burn down Chicago. I have encountered but a single case of legal entities being physical causes: as Figure 10 reports, nonownership of fowl is a pathogen of water-filled cysts.[49]

> Hens normally roost in their owner's hut, and the latter sweeps out their droppings in the morning. Sometimes a hen will decide to nest and lay in another hut. Then it is the custom to leave her to hatch her brood there, the bird's owner giving one of the hatched chickens to the owner of the hut which sheltered them. This gift is sanctioned by the belief that a man who sweeps up the droppings of another man's bird is exposed to the risk of developing a water-filled cyst unless one of the chickens can be called his own. Sweeping the droppings of one's own hen is not supposed to have any ill effects.

FIGURE 10
Forensic Medicine

Theoretical terms in the first instance *describe* the partition that the w_{ij}s make in the space of the hidden units. Imagine the training set is Llewellyn's casebook on Sales,[50] a venerated text by a legal realist. The traditional casebook is Langdell's:[51] "By present-day standards, it was a curious work. It contained only 336 cases, only slightly edited,

---

47. *Id.* at 92.
48. Holmes, *The Path of the Law*, 10 HARV. L. REV. 457, 459-60 (1897).
49. M. DOUGLAS. *Social and Religious Symbolism of the Lele*, in IMPLICIT MEANINGS: ES-SAYS IN ANTHROPOLOGY 9, 11 (1975).
50. K. LLEWELLYN. CASES AND MATERIALS ON THE LAW OF SALES (1930).
51. C. LANGDELL. CASES ON CONTRACTS (1st ed. 1870; 2d ed. 1879).

without commentary."[52]

Llewellyn's use of cases departed from tradition in several respects. The majority of 'cases' consisted of a summary of the facts together with the result, the reasoning being omitted entirely. Behind this were two ideas, inherited from Corbin and Cook; first the idea that at least as much significance should be attached to what judges do (the result based on the facts as the judges saw them) as to what they say by way of justification. Secondly, the idea that the facts of cases have a significance that transcends their significance as precedents and illustrations of doctrine: they are concrete illustrations of business situations, which give a flavor of practice beyond the particular legal issues involved; they are also excellent raw material for students to treat as problems, more closely related to 'real life' . . . .[53]

I display a representative case edited by Llewellyn as Figure 11.[54]

So Hart studies a set of these teaching cases, and bifurcates the space of the hidden units, liability (= mine) in one subspace, no liability (= rock) in the other. BC-11

> [Case No. 259] G. AMSINCK & CO. V. SPRINGFIELD GROCER CO., 300 Fed. 452 (S.D. Mo., 1924) the contract called for "Java white granulated sugar." There was evidence that sugar produced in Java is neither white nor granulated. S tendered Java sugar which B rejected, and S sued for the price. Held for B. The court seemed much influenced by the fact that the exact nature of "Java sugar" was unkown to all but a few experts and that B, a jobber of groceries, including sugar, probably relied on the word "granulated" in the contract.

FIGURE 11
A Case Stated Without *Essential* Theoretical Terms

---

52. Farnsworth, *Casebooks: Contracts Scholarship in the Age of the Anthology*, 85 MICH. L. REV. 1406, 1409 (1987).

53. W. TWINING, KARL LLEWELLYN AND THE REALIST MOVEMENT 132 (1973) (footnote omitted).

54. K. LLEWELLYN, *supra* note 50, at 312-13 (1930).

A note on calling black 'white'. Mitchell v. Henry, 15 Ch. D. 181 (C.A. 1880), also declined to do it. Still, Swift said of attorneys, "I said there was a Society of Men among us, bred up from their Youth in the Art of proving by Words, multiplied for the Purpose, that *White* is *Black* and *Black* is *White*." J. SWIFT, GULLIVER'S TRAVELS 295-96 (Crown ed. 1947). Now news from a sister discipline. *Cf.* T. GREEN, *Lectures on Logic*, in 2 THE WORKS OF THOMAS HILL GREEN: PHILOSOPHICAL WORKS 157, 161 (R. Nettleship ed. 1886). Theology teaches: "That we may be altogether of the same mind and in conformity with the Church herself, if she shall have defined anything to be black which to our eyes appears white, we ought in like manner to pronounce it black." R. POPKIN, THE HISTORY OF SKEPTICISM FROM ERASMUS TO SPINOZA 4 (1979) (quoting Saint Ignatus Loyola). Finally, to unite the disciplines, recall that in Rabelais, attorneys "tourne le noir en blanc," incited by Satan. F. RABELAIS, LE TIERS LIVRE 299 (M. Screech ed. 1964).

implicitly distinguishes metal things; Hart, contracts.

Why make the theories explicit? First, see that *any* output of BC-11 or Hart is verbal. They say 'Mine' or 'Liability', then get error messages. With explicit theories, the student's output can be *about the theories*, and itself subject to error messages. For instance, '<1, 0>' and '<0, 1>' might be read 'Metal' and 'Nonmetal' or 'Contract' and 'No contract, and Kingsfield respond, e.g., 'No, there is a contract in *Amsinck*, it is just that S is not liable for the price'. The anticipated gain is that BC-11 and Hart would learn about liability, which Holmes says is what counts, quicker or more accurately. Such tuition would complement, not replace, that about mines and liability. The input at the second stage would be the right response in the first stage, '<1, 0>' or '<0, 1>', depending on whether a contract exists, plus some description of the factors bearing on liability, given a contract or not.

The first, theoretical stage would elicit a space of hidden units like that of Figure 9, but with subspaces deciding whether there is a contract instead of whether there is liability. Hart partitions that space too, implicitly using a new theory. This new theory is for instance that contracts (old theoretical term) are promises (observational term) with consideration (new theoretical term).

And the process is repeated, with Kingsfield next assigning *Hamer v. Sidway*,[55] in which a nephew who abstained from many evil things until twenty-one gave consideration, and *Kirksey v. Kirksey*,[56] in which a sister-in-law who moved seventy miles to a new home did not. Holmes' 'bargain' must be the term for the new partition that plays the role 'contract' and 'consideration' do for the old: the uncle was bargaining, the brother-in-law was not. Here however contract law exhausts its theoretical language, cannot further describe its own theory, although the theory is still there. Description fails because 'bargain' is only an heuristic appeal to ordinary language, or because 'consideration' and 'bargain' are more nearly interdefinable than the second belonging to a theory about, being an analysis of, the first.

Three brief concluding remarks.

A. Once Kingsfield or Hart has w_{ij}s that only infrequently elicit error messages, he can apply his theory to correct the teaching set: reject some of its cases as being incorrectly decided.

---

55.   124 N.Y. 538, 27 N.E. 256 (1891).
56.   8 Ala. 131 (1845).

B. Leibniz wrote he had prepared

> a table, comparable in size to a map, which uses a unique arrangement and method to present the entire *common private law* of the Empire today, with all of its fundamental rules and propositions, and reduces them to *first principles* so that any one who understands this table, or has it lying before him, can decide any fact or case of private law, and at once put his finger on the basis for the decision in the table itself . . . .[57]

He did not dream of the w_{ij}s underneath such a table. If Hart's neurons can each have ten different outputs, there are $10^{100,000,000,000,000}$ possible theories. To get an idea of how many that is, reflect that the universe contains 10^{87} elementary particles.[58] There are that many possible pieces of language, one for each theory, but the descriptions of the theories become prohibitively long. If the neuronal outputs can take any real values between 0 and 1, there are uncountably many theories. Then because there are only countably many finite linguistic strings, the constraint becomes not physical but logical. Either way, almost all theories can be only incompletely expressed.

C. One can *buy* a PC program that simulates a neural network. The description entices, but the uses are frivolous:

> For example, harness racing data can be fed into the spreadsheet front end of *BrainMaker Professional* and the program can "learn" to find coincident data patterns that indicate a winning horse.
>
>   . . . .
>
> With *BrainMaker Professional*, you can feed in data from all the football games played this year, as well as the final scores. By the end of the season, the program should be able to guess the scores based on the pregame data.[59]

Fine, but maybe one can teach it law, too. The trick is to encode cases so the computer can understand them, yet not build a legal theory into the data, lest the computer adopt that theory, not construct its own.

---

57. Letter from Gottfried Leibniz to John Philip, March 27, 1699, in Loemker, *Introduction: Leibniz as Philosopher*, in G. LEIBNIZ, GOTTFRIED WILHELM LEIBNIZ: PHILOSOPHICAL PAPERS AND LETTERS 1, 6 (2d. ed 1969).

58. Churchland, *supra* note 27, at 94.

59. Dvorak, *BrainMaker Professional, Version 1.5*, PC MAG., January 15, 1991, at 168. A better program, in my opinion, is NeuralWare, Inc's *NeuralWare Professional II/PLUS*.