CATASTROPHE THEORY AND LAW

R. KEOWN

Department of Mathematics
University of Arkansas
Fayetteville, Arkansas 72701, USA

1. INTRODUCTION

This paper considers the utility of the cusp catastrophe as a model for display, analysis, and interpretation of a line of cases dealing with a well-defined legal issue. We give a very brief discussion of the judicial process in the United States, the form of presentation of decisions, and the methods of research common to lawyers in this country. The paper presents this information in view of the facts that the ordinary professional in mathematics or computing is unlikely to be aware of many of the details, while readers outside the United States are likely to be familiar with quite a different legal system. At the present time, research on the continent in this area apparently is somewhat in advance of that here [1].

Recent investigations of the American Bar Foundation have made consideration of the proposed application of catastrophy theory both possible and timely. Haar, Sawyer, and Cummings [2] analyzed 79 cases which came before the Supreme Court of Connecticut dealing with zoning amendment litigation over a period of more than twenty years, beginning in 1950, as a basis for their model. We recommend a study of their investigation to nonlawyers who may be or may become interested in the application of mathematics, statistics, and computers to the law. Their paper discusses the legal issues at some length in traditional lawyer fashion before describing the construction of a linear model based on their legal analysis.

A principal difficulty confronting the computer analysis of the law in the United States arises from the nature of the common law system and the manner of reporting the decisions, in that the courts do not decide cases in a fashion naturally adapted to legal analysis but rather chronologically as the parties bring their disputes before the tribunal. Consequently, the common law system of jurisprudence requires an elaborate finding system for its operation which, in the United States, has developed largely in private hands outside the judicial system proper. Two common procedures for locating the law pertinent to a given situation or fact pattern are either by means of legal digests or by means of legal encyclopedias. Entry through the digests is somewhat complicated, but entry through the encyclopedias parallels the use of ordinary encyclopedias.

We try to indicate in a very small space the problems of an attorney facing a client who brings a dispute for his advice or recommendation. Naturally, the dispute is apt to be one with which the attorney is familiar, but otherwise he is confronted with the problem of locating the law by means of the digests or the encyclopedias. This paper includes a bare bones discussion of the usual methods of legal research.

This information is presented in order not only to try to convince our readers that this application of catastrophe theory may survive the Sussman barrier, but also to introduce the reader to a very live area in which computers and mathematics have significant legal applications. As further evidence, we point out that there are two commercial legal data base systems on the market, WESTLAW, which is the product of the West Publishing

Company, the largest among those in the legal publishing field, and LEXIS, a service of Mead Data Central. There are two new journals: *Computer/Law Journal* published by the Center for Computer/Law, 675 South Westmorland Avenue, Los Angeles, California 90005 and *Rutgers Journal of Computers, Technology, and the Law*, Rutgers Law School, 15 Washington Street, Newark, New Jersey 07102.

2. THE LEGAL SYSTEM OF THE UNITED STATES

The judicial system of the United States began as and continues to be a part of the common law tradition. We use the term "common law" to distinguish the law of most English speaking countries from that of Roman law, modern civil law, and other systems of law. This English common law developed out of the custom of considering prior decisions of the courts as precedents by which to decide a case at bar in terms of examples or authority arising from the earlier holdings. Here, the fundamental concept is that of *stare decisis* which has been defined [3] as

". . . (T)hat when (a) court has once laid down a principle of law as applicable to a certain state of facts, it will adhere to that principle, and apply it to all future cases where the facts are substantially the same."

While the case law played the most important single role in the development of our legal system, it must be kept in mind that statutes, treaties, administrative decisions, and rules have enjoyed a significant part as well.

The judicial system of the United States is complex beyond that of most common law countries as a result of our federal system of government dividing the judicial and legal powers between the federal and the state governments according to constitutional dictates and long tradition. As a consequence of this division, our country has 51 judicial systems rather than only one as have most nations. Unfortunately, the division of power is not sufficiently sharp to permit one to search only the federal law or only the appropriate state law in deciding a legal question in all cases. This difficulty is complicated by an even worse one.

The sequence in which cases come before the U.S. and state courts for decision is completely independent of the subject matter of the suit or the legal issue in dispute, but depends entirely on the docket numbers assigned by the clerks of the various courts. The location of the case on the docket is a function of when the two attorneys can get their respective cases ready for trial and complete the required legal formalities. Thus, in 569 *Federal Reporter*, 2nd Series, there appears *Massachusetts General Hospital* v. *Weiner*, a case for declaratory relief and injunction, followed by *Norman* v. *Hall*, a criminal prosecution for rape and assault and battery, in turn succeeded by *de Walker* v. *Pueblo Intern., Inc.*, a suit for damages resulting from a false charge of shoplifting. The number 569 means, of course, that this is Vol. 569 of the *Federal Reporter*, 2nd Series, i.e., there are 568 earlier volumes of some 1000 pages each preceeding this one. West publishes 15 different reporter series of varying, but similar magnitude, which supposedly include all the case law of the 51 distinct jurisdictions of the United States. These cases are decided by the courts and appear in the reporters in the chronological sequence of their decision absolutely without regard for legal issues or subject matter content. The ordinary law library of a law school will contain thousands of volumes of case law, almost all reported in the manner outlined above.

These cases appear in a rather standard format with the title presented in the form *Jones* v. *Green* where usually, but not always, Jones is plaintiff and Green is defendant

in the suit. They appear frequently in an official reporter, but almost always in an unofficial reporter such as those published by West. For example, under *Rodriquez* v. *Taylor*, 569 F. 2d 1231 (1977), one finds docket number 76-2609-76-2641, decision by the United States Court of Appeals, Third Circuit, decided Dec. 27, 1977, followed by a short summary of the facts. Here, plaintiff brought suit challenging Philadelphia's policy of refusing to hire anyone over the age of 41 years as a security officer asserting that the policy was in violation of age discrimination in the Employment Act. Following this short summary of the case there is a sequence of headnotes, in this particular instance, there are 23, an unusually large number beginning with

1. Federal Courts (Key) 945, where in age discrimination in the Employment Act suit, the district court assessed liquidated damages in an amount equal to an underlying award of unpaid wages, insofar as the district court on remand might reduce the back pay award by setting off interim earnings, the equivalent liquidated damages penalty was to be likewise reduced. Age discrimination in the Employment Act of 1967, §§ 2–17, 29 U.S.C.A. §§ 621–634.

13. Federal Civil Procedure (Key) 2737. Although district courts have discretion in calculating an award of attorney's fees, exercise of such discretion must conform to proper standards and procedures; failure to adhere to generally applicable rules and criteria may constitute a "misuse" of discretion.

23. Federal Courts (Key) 945 . . .

The key numbers and headnotes are followed by lists of counsel appearing in the case then by the opinion rendered by the court. These key numbers and headnotes form the basis of one of our principal law finders, *The American Digest* of the West Publishing Company. When a West editor receives an opinion handed down by a court, he prepares the corresponding West reporter edition by carefully reading the case for its principal points of law and classifying each of them under the more than four hundred, including such items as automobiles, contracts, infants, wills, zoning, and other topics of the West digest system. He further searches the case for statutes construed, points involving the Federal Rules of Civil and Criminal Procedure, issues concerning the Federal Rules of Evidence, and words and phrases defined by courts in this case. West first publishes the edited version in one of its advance sheets which contains an index of the statutes construed, the issues involving federal rules and other items mentioned above at the front of the publication. In particular, the key numbers are collected for all the cases reported in the advance sheet and are presented alphabetically by topic, each topic containing its corresponding key numbers in numerical order. When the advance sheets are periodically accumulated and replaced by their corresponding bound volumes, these indices are similarly accumulated and placed at the front of the bound volumes. The collections from the bound volumes are accumulated in turn and compiled by West in *The American Digest* and in a companion publication *Words and Phrases*. We try to illustrate the use of the key numbers by a short introduction to legal analysis.

Suppose Mr. Speck appears in Mr. Land's law office with the story that he suffered an injured back at a college wrestling match when the referee was shoved from the ring and fell on Mr. Speck. Although Mr. Land has spent his prior career reading real estate abstracts, he must now elicit the crucial facts from Mr. Speck and determine the relevant law. Further discussion reveals that although Speck had purchased a reserve seat ticket, he was standing in an area forbidden to spectators when the accident occurred. One possible approach by Land to the relevant case law is through the *Descriptive Words* section of *The American Digest*. He enters by developing a list of words under the general headings of Parties, Places and Things, Basis of Action or Issue, Defense, and Relief Sought in the context of the facts and circumstances obtained from Mr. Speck.

Under Parties he may list spectator, wrestler, referee, promoter, and college with which to begin his search. Beneath the topic REFEREE, he finds

Wrestling match referee thrown from the ring and injuring spectator, contributory negligence. THEATER 6.

Beneath the topic SPECTATORS, he finds

Injuries—athletic events, assumption of risk. THEATERS 6(18).

Therefore, Land obtains the key numbers, THEATERS 6 and THEATERS 6(18). These key numbers furnish the subject matter entry into the case law. By going to them in *The American Digest*, Mr. Land can determine those cases which deal with the issues of Mr. Speck's injury and can determine whether or not Mr. Speck has a legal claim against the college. In addition to this procedure, West provides an entry into the key numbers by means of topic analysis and by means of a table of cases. With respect to these, if Mr. Land can remember the name of a plaintiff or defendant in a relevant case, then he can locate the case in the table of cases and thereby find the appropriate key numbers. Naturally there are many other approaches to the law and other publishing companies, in particular, The Lawyers Co-Operative Publishing Company which provide an approach to the case law. We turn to another procedure.

The West company publishes *Corpus Juris Secundum*, a legal encyclopedia, which not only illuminates the law in the manner of the usual encyclopedia, but also cites the reader to topic, key numbers, and cases. Consequently, it is the most logical first source when an attorney is confronted with an unfamiliar legal problem. Doubtless, Mr. Land would have been well-advised to begin here.

Before leaving this ridiculously short introduction to the methods of legal research, it should be noted that the LEXIS and WESTLAW legal data bases are arranged for searching according to these ideas. Such data bases help alleviate one of the chronic problems of the lawyer, to wit, that the law is in perpetual flux. Consequently, having found a case in point, the attorney must make every effort to determine if the case has been overturned or distinguished by a later opinion. Before the advent of these data bases, Shepard's Citations were the traditional method for locating later cases which dealt with the same point of law or legal issue.

3. THE LINEAR MODEL OF HAAR

Having outlined the judicial system of the United States and some of the basic tools of legal research, we wish to discuss a linear model of a specific area of law developed by C. M. Haar and others [2], but first we wish to quote from his paper:

"For Tiresias and the soothsayers of antiquity, prediction of Future events was an art clouded in mystery, wrapped in riddles. Today, lawyers advise their clients in carefully guarded professional terms about the probable future of potential litigation. Legal precedents hidden in the corners of recorded cases are scrutinized as closely as were the entrails of pigeons by the high priests of Apollo. Lawyers expend time and energy probing case law, searching for patterns, discerning trends, then struggling to impress the court with the similarities or differences between their clients' cases and the cherished legal precedents. Inferences are drawn from footnotes; asides are put under a magnifying lens; omissions, dissents, and changes in phraseology are all paraded out. Yet this process—in crucial aspects—is no less mysterious than that of the Greek seers and reflects as much intuition, resulting from personal immersion in the cases, as it does a rational process of analysis and dissection . . . Just as other areas of the social sciences have turned to computer analysis for assistance in sorting and analyzing complex sets of data, so could law."

Table 1.

Variable	Basic model	Model with Scale 9 added	Model without Scale 4
Var. 011	0.16510	0.16510	0.18324
Var. 012	−0.55106	−0.53905	−0.59198
Scale 7	−0.25548	−0.24747	−0.25450
Scale 8	0.21506	0.20939	0.21336
Scale 6	−0.19320	−0.18397	−0.19139
Scale 11	0.11884	0.10477	0.12014
Scale 10	0.05265	0.05629	0.06746
Scale 4	0.05769	0.05518	—
Scale 9	—	0.04473	—

Haar, Sawyer, and Cummings formed a team of two lawyers and one statistician to make a study of the case law developed by the Connecticut Supreme Court in zoning amendment cases. The two lawyers made a conventional legal analysis of 79 cases and determined a collection of 167 issues which possibly influenced the decisions. A significance test was run on this collection with the result that only 40 of the issues or variables were retained for analysis. With only 79 cases, the number 40 was too large for regression analysis so that a correlation matrix was computed for 32 of the variables, 8 having been eliminated by other arguments. The final result was a grouping of the variables into 11 scales which are listed below:

Scale 1—Compatibility Indicated by Change in the Character of the Neighborhood
Scale 2—Use not needed
Scale 3—Adequate Physical Planning
Scale 4—Public Interest Planning and Zoning Techniques
Scale 5—Compatibility from an Economic Perspective
Scale 6—Zone Change Detrimental
Scale 7—Physical Services Inadequate
Scale 8—Compatibility Indicated by Large Uniform Blocks
Scale 9—Good Planning Practice
Scale 10—Character of Area Supports Change
Scale 11—Large-Area Zoning.

In addition to these scales, two of the original variables were retained in the regression analysis, to wit,

011—The Court of Common Please Approved/Denied the Zone Change
012—The Zoning Authority Denied/Approved the Zone Change.

Variables 011 and 012 were not included in the factor analysis since the investors wished, for a variety of reasons, to include them in the regression model. The efforts of Haar and his colleagues lead to three different linear models, each of the general form

$$E = A + C_1 V_1 + \cdots + C_k V_k,$$

where the coefficients of the three models are listed in Table 1.

Using the conventional notion that plaintiff wins with a preponderance of the evidence in his favor, one may assume that if the value for a given fact pattern produces a number larger than 0.50, then there is a predition of verdict for plaintiff. the formula can be written for the basic model as

$$E = 0.56523 + 0.16121 V_{011} - 0.55106 V_{012} - 0.25548 S_7 + 0.21506 S_8 -$$

$$0.19320 S_6 + 0.11884 S_{11} + 0.05265 S_{10}.$$

Since the variables are assumed to take on only the values 0 and 1, those coefficients with plus signs are essentially for plaintiff and those with minus signs are essentially for defendant. Depending on the fact pattern, the court of common pleas will have approved $V_{011} = 1$, or will have disapproved $V_{011} = 0$; the zoning authority will have denied the zone change $V_{012} = 1$, or will have approved $V_{012} = 0$, and so forth and so on for the remaining.

Haar and his colleagues have assessed the interrelation of their linear model and conventional legal analysis [2] to which the reader is recommended for a fuller discussion. Since their observations probably apply generally to any successful mathematical or statistical model, we note some of them. Haar states that computer analysis leads to a better organized approach to a legal problem in that the data preparation requires a highly systematic examination of the cases. Furthermore, computer modelling displays the rule of the case as a relationship between the facts and outcome that can be mathematically expressed to the extent, of course, that the researcher can ascertain precisely what the facts were before the court. Moreover, the disciplined reading required for computer study of cases reduces the casual meandering through the key factors more likely in a study based on a case-by-case evaluation. As a final bonus, the evaluation of the research and conclusions by another lawyer not familiar with the cases is simplified by the organization necessary for the computer. Such presentation eases the task of a newcomer to the selected area of law by the systematization of the material. We turn to a catastrophic model related to this study.

4. THE CATASTROPHIC MODEL

We base our discussion on general concepts developed by Zeeman, in particular, on those of Zeeman [4] and others concerning institutional disturbances. Thus, we observe that the factors underlying a judicial decision may be classified either as (1) evidence supporting the position of plaintiff denoted by the symbol P, or (2) evidence supporting that of the defendant denoted by D. One must keep in mind, of course, that who is plaintiff and who is defendant may depend on which party wins the race to the court house, not on the nature of the dispute involved. Since who is defendant and who is plaintiff is somewhat arbitrary, this may require a switch in the axes of any given model, i.e., defendant assumes the role of plaintiff and plaintiff assumes the role of defendant. We recall that plaintiff wins his case provided the trier of fact, sometimes a judge or sometimes a jury, find a preponderance of the evidence is his favor. Regarding D and P as conflicting factors in a judicial process enjoying a suitably discontinuous behavior, we arrive by the Zeemanian process at

Hypothesis 1. The graph G of the judicial procedure is a cusp catastrophe with plaintiff's evidence P and defendant's evidence D as conflicting factors affecting the outcome.

As usual, Fig. 1 depicts the surface G as split into two parts, an upper part denoted as *judgment for the plaintiff* and lower part denoted as *judgment for the defendant*. This surface lies over the D,P-plane, the *control space* in the standard terminology, and consists of a single sheet for the most part. However, above a V-shaped region loosely centered between the P and D axes, the surface G consists of an upper sheet which is part of judgment for the plaintiff and a lower sheet which is part of judgment for the defendant together with a section joining these two.

The *standard model* of the cusp catastrophe is given by the equation

$$z^3 = y + xz,$$

where z is called the *response*, x the *splitting factor*, and y the *normal factor*. The
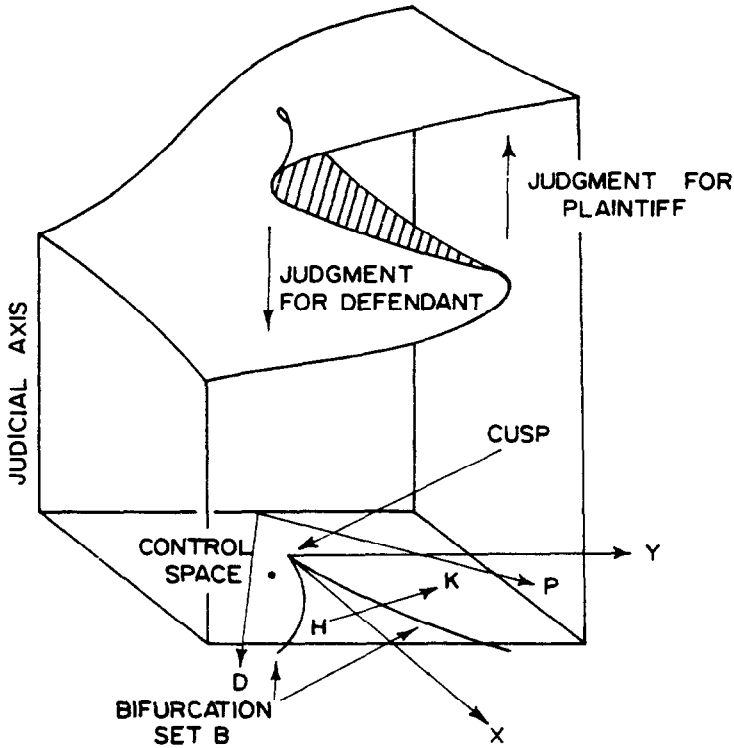
Fig. 1. The cusp catastrophe.

variables P and D are linearly related, as can be seen by Fig. 1, to the standard variables x and y. In terms of the standard variables, the graph of the surface G corresponds to the portion given by

$$3z^2 \geq x,$$

the threshold surface linking the upper and lower parts of G, judgment for the plaintiff with judgment for the defendant, corresponds to

$$3z^2 < x,$$

and the fold curve between them to $3z^2 = x$. The projection of the fold curve into the control space, now the x,y-plane, is given by

$$27y^2 = 4x^3$$

which defines the two sides of the V-curve known as the *bifurcation set B*. The bifurcation set B determines the catastrophic behavior of the judicial process. To see this, in Fig. 1 consider transversing the line HK in the direction from H to K and notice that as one crosses the right branch of the bifurcation set B there is a catastrophic jump from the judgment for defendant sheet to judgment for the plaintiff. Such sudden jumps are characteristic of the discontinuous properties of the catastrophic models, i.e., one gets a sudden change.

There are various reasons why small shifts occur in the structure of the catastrophic

surfaces. For example, society contains numerous natural plaintiff–defendant pairs including creditors v. debtors, mortgagees v. mortgagors, and insurers v. insureds along with many more who are eternally trying to better their posture before the courts. In view of this, all of them attempt to obtain favorable legislation, bring propitious rather than unpropitious cases before the courts and to write one-sided clauses favoring themselves into their contracts.

Hypothesis 2. There is a tendency for the judicial process as a whole to avoid the extremes of "judgment for the plaintiff" or "judgment for the defendant."

Zeeman regards this phenomena as a sort of flow on the catastrophic surface representing a feedback from the parties to the court which tends to prevent final stability in the system. There are other influences on the judicial process brought about by such things as one lawyer being more effective than another, one jury being more objective than another, one judge having more judicial competence than a brother with the overall result that in addition to feedback there is a certain amount of what Zeeman calls "noise" in the process. He visualizes such noise as forcing a particular case under consideration off the stable locus. As a consequence, there may occasionally arise catastrophic transfers from the mere presence of noise in the system. Zeeman presents this observation as a hypothesis.

Hypothesis 3. External events, or internal incidents, within the judicial process may be represented as stochastic noise.

Re-examining the model of Haar without Scale 4, we note that for plaintiff one has

Scale	8	0.213	Compatibility Indicated by Large Uniform Blocks
Var.	011	0.183	Court of Common Pleas Decision Favorable
Scale	11	0.120	Large Area Zoning
Scale	10	0.067	Character of Area Supports Change.

For defendant, Scale 6 consists of two variables and has a value of -0.19139. The two variables with their respective weights are

170	-0.11037	Adverse Impact on Adjacent Lots—Weight 1.80
106	-0.07358	Character of Area Improving—Weight 1.20.

With this splitting of Scale 6, one has for defendant

Var.	012	$-.59198$	Zoning Authority Decision Unfavorable
Scale	7	$-.25450$	Physical Services Inadequate
Var.	170	$-.11037$	Adverse Impact on Adjacent Lots
Var.	106	$-.07358$	Character of Area Improving.

Turning to the problem of reducing a nine-dimensional model to a three-dimensional one, we proceed in the following manner. For plaintiff assume there are but four issues which can be found in any fact pattern or that can be offered in evidence to the court, namely, S_8, V_{011}, S_{11}, and S_{10}. We denote the subset of issues present by means of a binary number, where 1 means that S_{10} is present alone, 11 means that S_{10} and S_{11} are present, 100 means that V_{011} is present alone, and 1000 means that S_8 is present alone. According to the scheme indicated, each of the integers 1 through 15 defines a unique fact pattern for plaintiff. Thus, the weight of evidence corresponding to the fact pattern defined by 111 equals $0.067 + 0.120 + 0.183 = 0.370$. (Note that one progresses from small to large probabilities with the patterns.) If one applies this device to all fact patterns for plaintiff and defendant, one obtains Table 2.

Table 2.

Number	Plaintiff	Defendant
1	0.067	−0.074
2	0.120	−0.110
3	0.187	−0.184
4	0.183	−0.254
5	0.250	−0.328
6	0.303	−0.364
7	0.370	−0.438
8	0.213	−0.592
9	0.280	−0.666
10	0.333	−0.702
11	0.400	−0.776
12	0.396	−0.846
13	0.463	−0.920
14	0.516	−0.956
15	0.583	−1.030

From their discussion, it appears that Zeeman and his co-workers made a linear analysis of their prison riot data before settling upon a method of computing tension and alienation somewhat analogously to the calculations reported in Table 2. After establishing these estimates, Zeeman made a time-series analysis of his data to locate the bifurcation set. We neither have the raw data collected by Haar to construct a time sequence nor do we believe that such a sequence is necessarily appropriate for the analysis of Haar's data. Yet, the rather marked changes in probabilities of Table 2 favoring the plaintiff between 7 and 8 reminds one of a catastrophic shift. One visualizes a lower sheet of judgment for the defendant determined by suitable values of D, while P takes on the values 1, 2, 3, 4, 8, 9, and 10 and an upper sheet of judgment for the plaintiff determined by suitable values of D when P takes on the values 5, 6, 7, 11, and 12.

A great deal of arbitrariness arises whenever one attempts to reduce higher dimensional data to lower, as in our example, where we try to reduce the sets of issues to a linear set by means of the binary definition of subset. Necessarily, we offer this procedure only as a possibility for reducing the dimension of the problem. Nevertheless, the fact remains that the legal profession speaks regularly of the evidence for the plaintiff and for the defendant: In civil cases one party wins by a *preponderance of the evidence*, in criminal cases the prosecutor must establish the guilt of accused *beyond a reasonable doubt*, a motion for summary judgment is overruled whenever the judge holds there is *substantial evidence* for the opposing party, a motion for a new trial is granted whenever the judge feels the verdict is *against the clear weight of the evidence*, and our judicial system depends for its daily operation on these rather *ad hoc* estimates of the judges.

We turn to a few remarks about the practical problems of using computers for the application of mathematical and statistical methods to problems in law. The reader should realize these remarks are only to direct attention to some relevant papers, not in any sense to outline or summarize their ideas and arguments.

5. BASIC PROBLEMS IN COMPUTER-ASSISTED LEGAL RESEARCH

For law firms, which can afford the service, a number of data bases are available of which the best known in the United States are LEXIS and WESTLAW. The implementation and cost of these services are reviewed by a dedicated user [5] who observes these particular services tend to be complementary rather than competitive. The principal difficulty with their use is the cost which tends to outrun the budget for smaller law firms.

While there has been an effort to provide such service by the use of public terminals [6], these have enjoyed limited success for a variety of reasons. Nevertheless, these data bases are commercially available and provide decided cases from the federal and state courts, various articles from the United States Code, and various federal regulations such as tax law and trade regulations.

The WESTLAW system retrieves information according to their key number system and LEXIS by means of words and phrases agumented by means of various Boolean Algebra operations. In either instance, the information is essentially the same as would be obtained by traditional legal research although it may be more complete and more quickly obtained.

In addition to the cost, a second difficulty is that the principal problem of the lawyer is the analysis, rather than the accumulation, of a large number of cases in point. Clearly the efforts of C. M. Haar to provide a linear model of a line of cases constitutes a nontrivial achievement in solving this problem. Not only should the results provide substantial aid to parties involved in zoning amendment cases in Connecticut, but their procedure obtains all that could be reasonably expected of mathematical analysis at the present time. Consequently, there is a need to apply their methodology to other well-defined legal situations perhaps by teams from the law and mathematical divisions of universities.

Unfortunately, the method of Haar requires a large component of traditional legal investigation, likely more than goes into many important legal cases, before computer analysis takes place. Consequently, their present approach may be better adapted to academic law review work than to law practice per se. A procedure is needed whereby the practicing attorney may start with a data base such as WESTLAW or LEXIS and produce a model, or a satisfactory substitute, similar to that of Haar, without the need of the protracted legal study.

A member of the Norwegian Research Center for Computers and Law has suggested a somewhat different approach to legal information retrieval [7]. Bing discusses the possibility of a retrieval strategy based on natural language inquiries rather than on some artificial system and compares the efficiency of the two. Bing and Harvold have won the Norwegian Royal Academic Gold Medal for their book, *Legal Decisions and Information Systems*, available in North America from Columbia University Press. Various articles [8, 9, 10] investigate other techniques of legal data retrieval and analysis.

So far, the author has located no articles dealing directly with applications of numerical taxonomy [11] or with cluster analysis [12] in the law. The methods of data retrieval employed by WESTLAW or LEXIS are unknown to the author, but assuming they use some vector-based retrieval strategy [7], as seems likely, both numerical taxonomy and cluster analysis could be employed. In a vector-based retrieval strategy, each document in the data base is assigned a name consisting of a vector whose length is that of the number of words (or key words) in the data base. For example, in cases dealing with motor vehicle accidents, one might obtain a segment of the document vectors consisting of

$$accident \ldots car \ldots negligence \ldots traffic\ light \ldots corner,$$

where in Case A the number 10 shows that the word *accident* occurs 10 times in the opinion, the number 5 implies that the word *car* appears 5 times in the opinion and so forth and so on. It seems virtually certain that a more complete analysis could be obtained by the further use of numerical taxonomy or cluster analysis on computers if the vectors defining the cases, or substantial segments thereof, were made available to the investigator.

REFERENCES

1. J. Bing and T. Harvold, *Legal Decisions and Information Systems*, Norwegian University Press, Oslo (1977).
2. C. M. Haar, J. P. Sawyer, Jr., and S. J. Cummings, Computer Power and Legal Reasoning: A Case Study of Judicial Decision Prediction in Zoning Amendment Cases. *A. B. F. Res. J.*, 651–768 (1977).
3. E. M. Wise, The Doctrine of Stare Decisis. *Wayne Law Rev.* **21**, 1043 (1975).
4. E. C. Zeeman, *Catastrophe Theory: Selected Papers* (1972–1977), Addison-Wesley, Reading, Mass. (1977).
5. R. M. McGonigal, Implementation and Cost Effectiveness of Computerized Legal Research—LEXIS and WESTLAW Compared. *Computer/Law J.* **1**, 359–378 (1978).
6. F. M. Greguras and L. L. Carlile, Report of the Public Terminal Study. *Computer/Law J.* **1**, 255–357 (1978).
7. J. Bing, Legal Information Retrieval Systems: The Need for and the Design of Extremely Simple Retrieval Strategies. *Computer/Law J.* **1**, 379–399 (1978).
8. R. C. Lawlor, Amplification Factor of an Information Retrieval System. *Jurimetrics J.* **15**, 289–302 (1975); The Chancellor's Foot: A Modern View. *Houston Law Rev.* **6**, 630–664 (1969).
9. L. T. McCarty, Reflections on Taxman: An Experiment in Artificial Intelligence and Legal Reasoning. *Harvard Law Rev.* **90**, 837 (1977).
10. J. A. Sprowl, Automating the Legal Reasoning Process: A Computer That Uses Regulations and Statutes to Draft Legal Documents. *A. B. F. Res. J.*, 1–81 (1979).
11. P. H. A. Sneath and R. R. Sokal, *Numerical Taxonomy*, Freeman, San Francisco (1973).
12. J. A. Hartigan, *Clustering Algorithms*, Wiley, New York (1975).