Artificial Intelligence and Law **9:** 165–188, 2001.

Evidence marshaling for imaginative fact investigation ⋆

DAVID A. SCHUM
School of Information Technology and Engineering, School of Law, George Mason University, USA

For the past ten years Peter Tillers and I have been investigating the process of discovery and the imaginative reasoning it involves. Though our work has been performed within the context of law, we believe it applies in other contexts as well. One basic premise upon which our work rests is: *How well we marshal or organize our existing thoughts and evidence influences how well we are able to generate new ideas in the form of hypotheses, new evidential tests of all hypotheses being considered, and defensible arguments linking our evidence and hypotheses.* Existing thoughts and evidence can be marshaled, combined, or juxtaposed in various ways to meet different requirements that arise as the process of discovery unfolds over time.

The work being described rests on other discovery-related studies in a variety of disciplines. As expected, we obtain different perspectives on discovery from different disciplines. One reason is that elements of the process of discovery are situation-dependent. In some contexts we may already have in existence an extensive base of information and wish to see what this information reveals. In other contexts, such as in fact investigation in law, we must often begin an episode of discovery with little or no base of relevant information. Further, in some situations we may have existing collections of data that may be analyzed statistically in various ways. But in other situations, such as in law, we usually encounter singular, unique, or one-of-a kind events for which no meaningful statistical analyses are possible. What does seem to be common in discovery in different situations is the need to examine different combinations of existing information. Different combinations or juxtapositions of existing information may, in different ways, suggest new ideas and new avenues of inquiry. In this paper I describe a prototype system that allows a person to juxtapose thoughts and evidence in different ways, each of which is helpful in suggesting new ideas, new evidence to gather, and new questions to ask.

⋆ The research upon which this paper is based was supported by the National Science Foundation under Grants SES-8704377 and SES-9007693 to George Mason University and by the Law and Economics Center at George Mason University. The author is most grateful for this support.

1. Studies of discovery and evidence marshaling

In a recent work in what is now called the *science of complexity*, Coveny and Highfield (1995, p. 279) report: a belief, held by many persons, that the human brain is the very *cathedral of complexity* in the known universe. Of all the "services" or activities taking place in this cathedral, none are more interesting, important, and complex than our ability to generate new ideas, in the form of hypotheses, and new forms of evidence. Thus it is not surprising that study of our imaginative reasoning and creative abilities has been so difficult. Over the past twenty centuries or so many persons have commented on these abilities and have characterized the process of discovery and imaginative reasoning in different ways.

One of two species of inductive reasoning Aristotle mentioned in his *Posterior Analytics* involves intellectual *intuition* [Greek word *nous*], by means of which we guess at explanations based upon our perceptions of relations among things (Kneale 1952, pp. 24–37; Cohen and Nagle 1934, p. 275). Centuries later, Galileo supposed that we "reason backwards" in imagining causes from observed effects (Oldroyd 1986, pp. 55–56). Closer to our own time Sir Arthur Conan Doyle has Sherlock Holmes, in the case: *A Study in Scarlet*, explain to Dr. Watson that his [Holmes's] feats of discovery rested on his ability to "reason backward" from observed events to possible hypotheses, causes, or explanations [Baring Gould 1967, Vol. 1, p. 231]. At the same time Conan Doyle was providing Sherlock Holmes with his remarkable discovery-related abilities, Charles S. Peirce in the United States was discussing a new form of discovery-related reasoning he variously called *abduction*, *retroduction*, or simply *hypothesis* (1898, 1901, 1903). The semioticians Eco and Sebeok (1983) have provided a very interesting account of the similarities between Holmes's "backward reasoning" and Peirce's abductive reasoning. According to Peirce, we often encounter anomalous evidence we cannot explain by means of existing hypotheses. However, after a time we may often experience flashes of insight in which a new hypothesis occurs to us that explains this anomaly and, possibly, other evidence we may have. Peirce associated these flashes of insight with abductive reasoning.

Until quite recently, philosophers and logicians have been quite content to relegate the study of discovery and imaginative reasoning to others, such as psychologists, on grounds that discovery-related processes seem to follow no particular logic (e.g., Reichenbach 1968, p. 231; Popper 1968, p. 31). A question frequently asked is: How does anyone perform a logical analysis of the flashes of insight associated by Peirce with discovery? Peirce's term *abduction* is very much alive these days, particularly in the field of artificial intelligence. But in this field the term *abduction* has taken on a variety of different meanings. In fact, this term now occupies the status of a "wild card" and is used quite differently by different persons. For example, in some works *abduction* is taken to mean *inference to the best explanation* (Peng and Reggia 1987; Neapolitan 1990). This view has been criticised on grounds that what constitutes the "best explanation" is rarely obvious (e.g., Michalski 1994, pp. 19–20). In addition, this view mixes together the processes of discovery and justification. But Peirce himself allowed that abduction and induction are often combined (1901, pp. 152–153). Some persons in artificial intelligence interpret abduction to be the entire process of generating, criticizing, and accepting explanatory hypotheses (Josephson and Josephson 1994). Others have made distinctions among several different species of abductive reasoning (Eco 1983, pp. 198–220; Thagard 1993, pp. 52–65).

Of particular interest in our work are collected insights about the importance of marshaling or combining our thoughts and our evidence in the process of discovery. Mathematicians have taken a special interest in the combination of ideas in discovery. Poincaré (1908, pp. 2042–2045) noted that facts worthy of our consideration are those which reveal unsuspected kinships with other facts. But he also noted that only a small number of combinations of ideas will be fruitful and that our choice of which ones to investigate is all-important. Jacques Hadamard (1954, pp. 29–31) noted that discovery in mathematics, or anywhere else, takes place by combining ideas. What is quite interesting is that both Hadamard and Poincaré speculated that our subconscious mind goes through all combinations of stored ideas on a certain topic. When it finds an esthetically interesting combination, the subconscious mind then reports this combination to our conscious mind. This view is reflected in the more recent work of Roger Penrose (1991, p. 423) and also brings to mind the work of the neurophysiologist J. C. Eccles on neural correlates of discovery (1970, pp. 127–129). Albert Einstein also commented on discovery and the combination of ideas arguing that this combinatory play is an essential feature of productive thought (1954, pp. 142–143).

Others besides mathematicians and physicists have emphasized the importance of combining ideas. Another term C. S. Peirce used in connection with discovery is *colligation*, the process of combining ideas (Tursman 1987, p. 19). Peirce argued that the binding together or colligation of ideas occurs throughout the process of discovery. In his celebrated work: *The Act of Creation*, Arthur Koestler described a process he called *bisociation* (1989, pp. 44–45). In this process we bring together different frames of reference or perspectives. This combining of ideas often occurs as a result of important questions we ask. As he stated: "The bisociative act connects previously unconnected matrices of experience. ... Spontaneous flashes of insight occur in which familiar events are seen in a new light" (1989, p. 45). The semiotician and novelist Umberto Eco appreciated the fact that isolated thoughts or evidence items seem unimportant until we bring them together. In his novel *Foucault's Pendulum* (1988, p. 225) he has a major character say: "No piece of information is superior to any other. Power lies in having them all on file and then finding the connections. There are always connections; you have only to want to find them".

No person ever devoted more attention to the marshaling of thoughts and evidence than the American jurist John H. Wigmore (1863–1943). Wigmore developed analytic and synthetic methods for drawing conclusions from masses of evidence.

Wigmore's work forms the very first systematic study of what we now term *inference networks* (Wigmore 1913, 1937). In addition to being interested in the analysis and synthesis of existing masses of evidence, Wigmore suggested several devices for organizing thoughts and evidence during the process of discovery (1937, pp. 994–1003). As we have noted elsewhere, our own work rests heavily on Wigmore's (Tillers and Schum 1988).

2. A focus on discovery problems in law

To set the stage for discussion of a prototype system for marshaling thoughts and evidence, I mention various characteristics of the process of discovery as it occurs in the field of law. Attention to these elements allows me to contrast our work on discovery with that now in progress in other disciplines. The term *discovery* has at least two interpretations in law. It is used with reference to the rules allowing one party in a contentious matter to obtain certain evidence held by the opposing party. We have termed this form of discovery *legal discovery.* However, our interest has centered on another form of discovery we have termed *investigative discovery* (Tillers and Schum 1991, pp. 939–941). By this term we refer to discovery-related activities that occur during the process of *fact investigation* in which attorneys and investigators seek to generate their own *hypotheses* concerning possible charges or claims, their own *evidence* bearing on these charges/claims, and their own *arguments* in support of the relevance, credibility, and probative force of this evidence. We have supposed that investigative discovery begins prior to legal discovery and that eventually the two may overlap as a trial or other form of settlement becomes imminent.

The process of investigative discovery unfolds over time. As it unfolds, we have evidence in search of hypotheses at the same time we have hypotheses in search of evidence. Every episode of investigative discovery is unique, as any attorney or investigator can testify. There are several important dimensions to this uniqueness. In some situations, hypotheses regarding possible specific claims or charges emerge rapidly. For example, an investigator might find a fingerprint or a blood sample linking a specific person to a crime. In other cases, however, initial hypotheses may be quite vague, such as: "The killer was a male". Only after extensive investigation are more refined or specific hypotheses generated. This same characteristic is evident in both criminal and civil fact investigation.

A second characteristic concerns what I will term the *initial conditions* of an episode of investigative discovery. By this term I mean the amount of available information an investigator has to go on when a case begins. It is certainly true that attorneys and investigators bring their own accumulated knowledge and experience to bear on each case. But the trouble is that, in most instances, an episode of fact investigation begins with very little or no *specific* information about the case at hand. Where does such information come from? The answer, of course, is: *By asking questions.* This fact exposes the second major premise upon which our work has rested: *Well-developed search heuristics are a necessary but not a sufficient condition for productive and efficient investigative discovery.* At the beginning of an episode of investigative discovery in law, and in other fields as well, there may be very little or no existing base of information to search. *What is also crucial in investigative discovery is the process of inquiry.*

As we begin the process of inquiry we recognize that the answers to some questions we ask will seem to lead us nowhere. The trouble is that we cannot always appreciate the significance of an item of information until we obtain other items. In many cases an item of information dismissed as irrelevant at one stage of an investigation becomes very relevant at another. We would all like very much to be able to ask the "right" questions. The trouble is that we may have no basis for asking the "right" questions until we have asked and answered other questions that arise as investigative discovery lurches forward. *A major objective in our work on the marshaling of thought and evidence has been to stimulate the process of inquiry.* We have been influenced by the work of Jaakko Hintikka and his colleagues regarding the importance of inquiry during discovery (1983a, b). Hintikka has argued that Sherlock Holmes's extraordinary discovery feats had nothing to do with abductive reasoning, but were possible because Holmes was so adept at asking "strategically important questions". Such questions generate new productive lines of inquiry.

Another major characteristic of fact investigation in law is that the events of concern are usually unique, singular, or one-of-a-kind. We cannot play the world over again a thousand times in order to determine the number of occasions on which a particular defendant breaches a duty of care in a negligence case or a particular witness in a murder trial gives untruthful testimony. Thus, for many of the events of concern in legal confrontations there is never any base of relevant evidence that can be analysed statistically. We do acknowledge, however, that statistical evidence has come, by degrees, to have greater importance in certain kinds of trials. However, its admissibility is frequently challenged. We became involved as court-appointed witnesses, in a recent dispute about the admissibility of statistical evidence in a narcotics sentencing decision (United States v. Shonubi, 998 F.2d 84 (2d Cir. 1993)) In this case the Court of Appeals unfavorable rulings concerning the use of statistical evidence, as specific evidence against a defendant, have been the subject of considerable recent debate (e.g., see Carson 1996, entire issue).

The final point I mention about investigative discovery is relevant to other contexts besides law. Discovery takes time and costs money and there are always decisions to be made about what "investigative tracks" to follow. Such decisions are very difficult since, absent clairvoyance, there is no way of determining in advance the consequences of following a particular track. Stated another way, when we decide upon a course of inquiry, we will never be certain about what answers we will obtain to questions we intend to ask. Nor is there any guarantee that we will receive any answers when these questions are asked.

Having mentioned some of the special attributes of investigative discovery in law, I can now begin to relate our studies of evidence marshaling and discovery to other discovery-related work performed by our colleagues in computer science and artificial intelligence. Over twenty years ago Herbert A Simon argued that a logic of discovery was possible and that what we call imaginative or creative processes can be captured in computational terms (1973, pp. 471–480). One of the best-known attempts to achieve computer-based discovery, at least in science, is the work of Langley et al. (1987). They viewed discovery as just another species of problem-solving and argued that heuristic search paradigms form the core of any species of problem-solving. In our view, stated above, search is a necessary but insufficient condition for productive and efficient discovery in many contexts such as law. The process of inquiry is at least as important. The work of Langley et al. on discovery has been received enthusiastically by some (e.g., Boden 1991), but not so enthusiastically by others (e.g., Holland et al. 1989; Wolpert 1992, pp. 64–67; Miller 1996, pp. 326–332). Miller argues that the computer programs Langley et al. employed already contained the results to be discovered. Wolpert argues that the work of Langley et al. demonstrates the wisdom of hindsight rather than the generation of new ideas. Wolpert also argued that a problem-solving metaphor for discovery is inadequate. I agree with Wolpert for reasons I have stated elsewhere (Schum 1994, pp. 469–470).

Our work on evidence marshaling may also be compared to the work now in progress on *knowledge discovery in data bases* [KDD] and a step in this discovery process called *data mining.* As stated recently (Fayyad et al. 1996, p. 6), KDD is the process of identifying "valid, novel, potentially useful, and ultimately understandable patterns in data". Patterns in data are identified using data mining techniques that commonly result in algorithms or models. These models are typically generated from data during the data mining process. This process rests on a variety of concepts from statistics, machine learning, and pattern recognition. The entire KDD process involves the evaluation and interpretation of patterns mined from data. As noted in the *Price Waterhouse Technology Forecast for 1996*, "... the capacity to look at the world awash in information and find significant patterns may very well be the best definition of an intelligent system" (1996, p. 640).

In many business, scientific, military, and other domains, very large data bases are maintained in which many data or records are available in a large number of fields, variables, or attributes. For example, a company that issues credit cards may have a variety of items of information on record for each of their customers. If this company has, say, 100,000 customers and has 100 items of information on each customer, it has ten million records in its data base. Data bases having hundreds of fields or variables and millions of records in each field are becoming more common. The trouble, of course, is to try to discover what *knowledge* is lurking in these large bases of information. Examined one or two at a time, the fields or variables provide only the most obvious knowledge. Many subtleties or complexities may be revealed if larger numbers of fields or variables could be examined jointly.

Stated differently, the discovery of novel and interesting patterns in data requires examination of combinations of more than two fields. Here is where trouble resides in KDD and data mining work as well as in our work on evidence marshaling.

As fact investigation proceeds, we begin to obtain answers to questions we ask. These answers, or items of information, begin to mount up rapidly and there quickly becomes no hope of examining all combinations or juxtapositions of these items, even if it made sense to do so. Looking through everything in the hope of finding something is not a sensible strategy. In short, we encounter the same combinatorial explosion in fact investigation as in the work on KDD and data mining. In both situations we need procedures for deciding which combinations or patterns of data will be most productive in supplying new knowledge. In our legal work, this new knowledge exists in the form of new hypotheses, new evidence, and arguments linking hypotheses and evidence. In other words, we need good heuristic strategies for search in fact investigation as we do in other contexts. As I now describe our work on evidence marshaling in law, additional parallels can be observed between this work and work on KDD and data mining.

3. MARSHALPLAN: A network of marshaling operations

The prototype computer-based system we have designed is called *MARSHALPLAN*. In other works we have given a more complete account of the steps in our research on evidence marshaling that influenced the design of this system (Tillers and Schum 1991; Schum and Tillers 1991). Our work on evidence marshaling is still in progress and we are investigating the use of *MARSHALPLAN* in other contexts such as intelligence analysis.[1]

When we began our work on *MARSHALPLAN* we had a fairly clear idea of what any attorney or fact investigator would like to have in a system for marshaling thoughts and evidence. At the same time, however, we recognized that systems having the following desirable capability are not likely to be available any time soon. Suppose, at some stage in fact investigation, an attorney or investigator queries a computer asking it to retrieve a combination of data A, B, and C that appears interesting. The computer responds by saying: "You are not asking the right question here. You would do much better to examine data A, F, and G together since they suggest Possibility Q, which you have not yet considered". A computer system having this somewhat impertinent capability would certainly be helpful, provided that the computer was able to generate important or productive new lines of inquiry. But we did suppose that having access to a computer system that allows thoughts and

[1] Peter Tillers and I collaborated in all phases of the design and development of *Marshalplan.* We are grateful for the encouragement and advice of Tod Levitt, President, *Information Extraction and Transport Inc.* (East Setauket, NY), as we are now attempting to go beyond the prototype stage of our evidence marshaling system. We also wish to record our thanks to Dr. Kenneth Auerbach, a criminal defense attorney in Silver Spring, Maryland, who helped us test the reality of our ideas about evidence marshaling in the crucible of day-to-day law practice.

evidence to be marshaled in various ways would be both feasible and heuristically valuable in assisting the user to generate new questions or new lines of inquiry. The marshaling operations we describe are heuristically valuable in different ways.

Shown in Figure 1 are fifteen different marshaling operations that form the basis of *MarshalPlan*. Each marshaling operation is indicated by an icon representing a stack of cards. This is a natural icon to use since our prototype system was constructed using the Hypercard™ system [Claris Corporation]. Each different marshaling operation is in fact a stack of cards whose backgrounds, fields, and buttons are made specific to particular marshaling operations. Though this system is in fact a network of marshaling operations, we have not shown any links or associations among the operations shown in Figure 1. One reason is that a user can, using appropriate buttons in a stack, go easily from one marshaling operation to any of the others. In short, the marshaling operations are all linked together.

The second reason for not showing any links is that we do not wish to convey the impression that there is either a natural or an enforced sequence in which the user employs these marshaling operations. As we noted, every episode of fact investigation is unique. The thoughts and evidence an investigator might have at the start of one episode might be entirely different from what this person has at the start of another. For example, in some fact investigations a user might immediately perceive a possible charge or complaint. In others it may be quite some time before definite possibilities are generated. Some fact investigations begin with tangible evidence; others begin with testimonial assertions. Finally, our system was designed for use by both parties in a dispute. Attorneys and investigators for the defense might make use of different combinations of these marshaling operations than attorneys and investigators for plaintiffs.

For convenience in indicating how many stacks a user might actually employ in an episode of fact investigation we show the cardinality of each stack. Observe that there is one of three symbols above each stack or marshaling operation: 1 = "exactly one"; 1... = "one or more"; and * = "many". The cardinality symbols we have chosen are those employed in current versions of the *Unified Modeling Language* [UML] as described in the *Platinum Paradigm Plus*™ *Methods Manual* (1996, 4-1-96). Some episodes of fact investigation will require just a single stack of the same variety, such as the one shown as *Chronologies.* Others require one or more stacks of the same variety such as the one shown as *Possibilities.* Still others will require many stacks of the same variety such as the Testimony stack. In a moment will describe each different stack and its cardinality.

Also observe in Figure 1 that I have shown the evidence marshaling operations as being in one of five *tiers.* The arrangement of these marshaling operations reflects the fact that different kinds of marshaling operations are required at different stages of fact investigation. For example, the *Argument Construction* operation cannot begin without the *Case Theory* operation which, in turn, cannot begin without the *Possibilities* operation. For reasons I explain later, there is one set of three marshaling operations called *Compilations* that lies outside this tier arrangement.

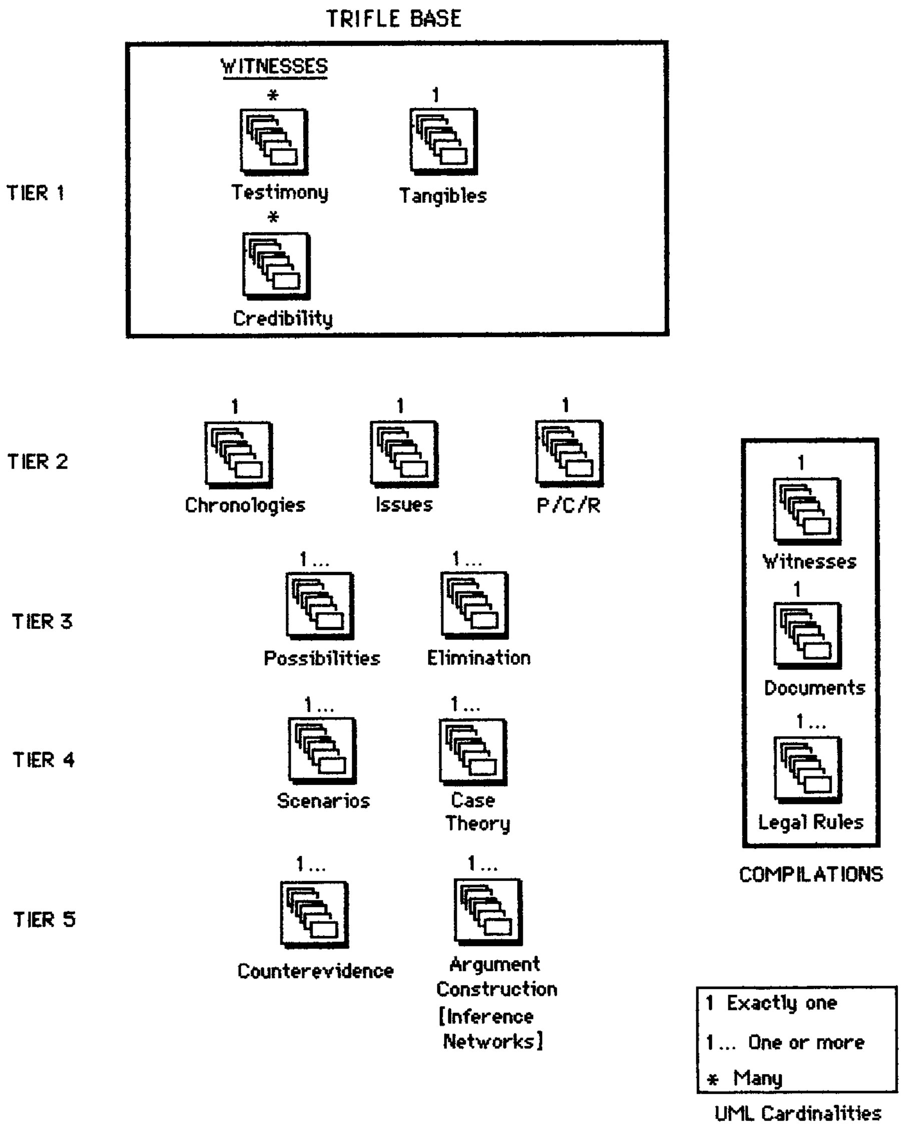

Figure 1.

Before I describe each marshaling operation, it is important to emphasize that our system provides no computational models or algorithms such as those that arise in *data mining.* Our system does allow for the generation of models in the form of inference networks, but there are no algorithms that describe these models.

I have described our system as one that allows the marshaling of thoughts as well as evidence. The thoughts we have often come in the form of questions we ask as investigative discovery proceeds. A most important element of MARSHALPLAN is that every marshaling operation or stack allows for the recording of questions we ask during the entire process of discovery. Keeping track of what questions we have asked turns Out to be a most important element of discovery as we mention again in Section 3.3.

3.1. MARSHALING TIER 1: A TRIFLE BASE

The three marshaling operations at Tier 1 we refer to as a *Trifle Base.* This term comes from an assertion Sherlock Holmes made in the case: *The Boscombe Valley Mystery.* In this mystery Homes explains to Dr. Watson how he was able to solve the case. Holmes tells Watson: "You know my method. It is founded on the observance of trifles" (Baring-Gould 1967., Vol. 2, p. 148). The *trifles* of concern in any episode of fact investigation are *details* that are observed during the investigation. Some trifles or details are found in testimonial assertions made by potential witnesses. Such assertions record the results of observations allegedly made by potential witnesses. In our system we have employed a separate stack for each witness. This accounts for the "many" cardinality above the stack labeled *Testimony.* Different cards in any witness's stack record different items of testimony. One trouble is that an assertion made by a witness, when carefully parsed, may contain many individual trifles or details. The user of our system is left to decide how many related details should go on a single card.

But there is an equally important class of details associated with testimony that merits special attention; this class includes evidence we may have *about* a witness and her/his credibility and competence. Such evidence is vital should the matter being investigated actually come to trial. The credibility and competence of any witness may be severely challenged. For each witness we have a separate stack that involves the marshaling of credibility-related evidence. In these *Credibility* stacks we marshal evidence about a witness in terms of its bearing on a witness's veracity, objectivity, observational sensitivity, and competence. The necessity for concern about these credibility-related attributes for testimonial evidence is discussed in detail elsewhere (Schum 1989; 1992; 1994, pp. 109–114, pp. 324–333). This credibility-related evidence is one specific of a form of evidence said to be *ancillary* in nature. Ancillary evidence is evidence about other evidence and its probative strength or weakness. Since there is one *Credibility* stack for each witness in an investigation, the cardinality of this stack will be the same as the cardinality for the *Testimony* stack.

Other trifles or details come in the form of various kinds of tangibles that may be gathered during an investigation. These tangibles can be objects, sensor records or images of various kinds, documents, tabled measurements, and charts. In our system we have employed a single stack, labeled *Tangibles* in Figure 1. Each tangible has a separate card in this stack. On this card appears a description of the tangible itself as well as other information bearing on the credibility and authenticity of this tangible. Such information concerns records of the chain of custody of the tangible item as well as other evidence that may bear on its accuracy, reliability, and authenticity. Such information is crucial in establishing the credibility of sensor records or images such as photographs, X-ray images, and police radar reports.

So, at the Tier 1 marshaling operations in our system we have the essential base of trifles or details upon which all other marshaling operations will depend. As fact investigation unfolds, the number of trifles we accumulate will grow very rapidly. It is here that the need arises for some heuristic strategies for examining combinations of these trifles in an effort to generate possibilities in the form of preliminary hypotheses about possible charges or complaints. As noted, in some cases we may get lucky and be able to generate a possibility or hypothesis directly from an individual trifle such as a fingerprint. In many cases, however, possibilities began to emerge only when trifles are examined in various combinations. This is the message conveyed by Poincaré, Hadamard, Einstein, Peirce, Koestler, Eco, Wigmore, and others I mentioned in Section 1. What we would like to have are metaphoric magnets or nets for attracting or capturing interesting and productive combinations of details from our trifle base. The following tiers of marshaling operations in MARSHALPLAN provide various magnets or nets for attracting or capturing trifle combinations. The next tier of marshaling operations allows the user to begin the process of examining various trifle combinations.

3.2. MARSHALING TIER 2: FORMING INITIAL TRIFLE COMBINATIONS

Perhaps the most obvious way of examining trifle combinations is by means of chronological orderings of events reported in trifles. This assumes that event times are reported in testimony and that tangibles are time-stamped. However, the actual time of occurrence of an event can only be inferred since witnesses may not be perfectly credible and tangible evidence may not be authentic. A witness might be untruthful or mistaken about the time of occurrence of some event or a photograph might be mislabeled as to the lime it was taken. Inferred temporal orderings of events can be quite useful heuristically in generating hypotheses with causal elements. As obvious as chronologies are as marshaling devices, there are several problems with chronologies that we discuss in more detail elsewhere (Tillers and Schum 1991, pp. 994–1000).

First, a single chronological ordering of all events reported or time-stamped in trifles can get very cluttered and be quite uninformative. In most legal "dramas" there are several or many "actors". Our system allows the user to construct separate event chronologies for each actor. When these actor chronologies are examined side by side, they can often reveal interesting possibilities that may be lost in the clutter of single or overall chronologies. Another chronology problem concerns the *granularity* of reported event times. Fact investigations in law involve many different kinds of events and situations. In some situations a temporal ordering of events with time intervals of days or longer is quite sufficient. But in other situations, temporal orderings of events in time intervals as short as minutes or seconds might be necessary. Our system provides for what we have termed a *tem-*

poral microscope, by means of which we can magnify or expand very short time intervals. This allows for temporal orderings of events in terms of minutes, seconds, or even shorter time intervals. The *Chronologies* marshaling operation or stack in Figure 1 carries cardinality 1 [exactly one]. Any episode of fact investigation needs only one such stack since the different cards in this stack can contain either the actor chronologies or the magnified chronologies just mentioned.

Now consider the marshaling operation called *Issues* in Figure 1. Here is a marshaling operation with distinct heuristic potential. As fact investigation proceeds, trifles are observed or collected and questions about them are raised; here is an example. A man is found dead on the floor of his garage. That he died as a result of a criminal act is indicated by the extensive nature of the head injury he suffered and by the absence of any indication that this injury was the result of an accident. I pause here for a moment to note that our system encourages the recording of trifles in the form of *negative* evidence, that which records the *nonoccurrence* of events. It also encourages the recording of *missing* evidence, evidence we search for but cannot find or evidence we request that is not produced. As we all know, negative evidence and missing evidence are not the same; evidence of absence is not the same as absence of evidence.

Returning to the man in the garage, suppose a small packet of white powder is found in his shirt pocket. Later analysis reveals that this packet contains uncut cocaine. The question immediately arises: Was this man killed during a drug deal that went sour? The issue raised by this question now serves as a magnet for attracting other trifles we now have, or may later gather. During the investigation, other trifles are observed, new questions are asked, and new issues are raised. Each issue raised can serve as a magnet for attracting new combinations of trifles already recorded and can, of course, suggest new lines of inquiry. Our system provides one *Issue* stack, on each card of which is a different issue together with places for recording the existing trifles this issue attracted. Taken in combination, the trifles recorded under some issue may suggest a new possibility or hypothesis.

The third marshaling operation at Tier 2 in our system is labeled *P/C/R*, shorthand for the words *prospectant*, *concomitant*, and *retrospectant*. These terms are Wigmore's and refer to evidence about events that happen before, during, and after some target event(s) of interest in law (Wigmore 1937, pp. 994–1003). Recording trifles that occur before, during, and after some target event seems just another form of chronological ordering. However, Wigmore goes much farther by first suggesting, at least in criminal cases, specific kinds of evidence in each of these three classes. Prospectant evidence concerns such matters as character, motives, intentions, habits and customs. Concomitant evidence concerns events bearing on opportunity and means for committing a crime. Retrospectant evidence concerns various kinds of trace evidence that can be either physical [such as fingerprints, footprints, glass shards] or mental [behavior indicating consciousness of guilt]. Wigmore also suggested specific questions to ask of the evidential trifles in each of these three classes. Our P/C/R marshaling operation captures Wigmore's thoughts on this form of marshaling our thoughts and evidence. Wigmore provides us with another heuristically valuable form of marshaling to be mentioned at Tier 5.

3.3. MARSHALING TIER 3: POSSIBILITIES AND THEIR ELIMINATION

The marshaling operations at Tier 2 in our system have the major objective of assisting the user to generate hypotheses about possible charges or complaints in a case being investigated. Again, in some instances a single trifle may suggest a possibility or hypothesis. But more commonly the generation of possibilities rests on combinations of trifles. As I noted, in many cases initial hypotheses are vague, imprecise, or undifferentiated. For example, based on some combination of trifles we might entertain the possibility that the man found on the floor of his garage was killed by someone he knew. As new trifles are gathered, we may *refine* this hypothesis to read: This man was killed by someone he knew who was also a heavy-set white male in his late twenties This process of refinement involves incorporating additional details into the statement of a hypothesis. Eventually, hypotheses may become very detailed and identify particular persons and specific actions they may have taken.

As fact investigation proceeds, we may have many possibilities or hypotheses at different levels of refinement; some may be more specific than others. Our system allows for the marshaling of favoring and disfavoring evidence on each hypothesis being entertained. Different hypotheses require different stacks, but refinements of the same hypotheses can be made in a single stack as we obtain new trifles that bear on them. In addition, it may become apparent that there is an evidential basis for two or more charges or complaints against the same person(s). For example, a person might be charged with murder and with possession of narcotics with intent to sell. In civil matters, there may be more than one possible complaint against an individual or organization. The marshaling operation labeled *Possibilities* in Figure 1 has cardinality "one or more" since there may be several hypotheses being entertained and different charges or complaints being considered.

Our *Possibilities* marshaling operation or stack shown in Figure 1 has another characteristic I believe to be very useful. This operation, in addition to others to be mentioned. allows for the maintenance of what we have termed *intellectual audit trails.* Such audit trails record what questions were asked at what times, what possibilities were being entertained at various stages of an investigation, and what was the existing evidential base for entertaining these possibilities at various times. Suppose an episode of fact investigation seems to be going nowhere or, worse, it is shown later at trial that charges in a criminal offense or complaints in a civil case cannot be sustained. During post-mortem analyses of some alleged miscarriage of fact investigation, hindsight critics may emerge and tell the investigators what possibilities they "should have" been entertaining and when they should have been entertained. Carefully maintained audit trails during discovery may defuse hindsight critics who, on existing evidence, may not have been able to do any better.

Such audit trails may also be useful in capturing episodes of fact investigation that are later judged to be exemplary, possibly having pedagogical value in the training of investigators.

As fact investigation proceeds, some possibilities are naturally eliminated. The elimination process itself calls for careful judgments. The other marshaling operation at Tier 3 in Figure 1 carries the label *Elimination* with cardinality "one or more". Our studies of evidence marshaling have been influenced by the work of L. Jonathan Cohen on eliminative and variative induction (e.g., 1977, 1989). One of Cohen's major concerns is the *completeness* or *sufficiency* of the evidence we have for a conclusion. The more relevant questions that remain unanswered by evidence, the less confidence we can have in conclusions we reach. By the same token, when we eliminate some plausible possibility or hypothesis, we should have adequate evidential grounds for doing so. Nothing is more embarrassing in episodes of discovery than the elimination of some possibility that turns out, on hindsight at least, to be one we should have retained. Our Elimination operation simply allows the user to marshal specific evidence-based reasons why certain possibilities or hypotheses were eliminated. This affords additional protection against hindsight critics since, in light of this same evidence, the critic might have eliminated these possibilities as well. It may also happen that in recording reasons why we are considering eliminating some hypothesis, we may also be led to the generation of reasons why we should still keep it alive. Perhaps there are relevant questions we have not yet asked that concern this hypothesis.

3.4. MARSHALING TIER 4: STORIES, SCENARIOS, AND CASE THEORIES

As a possibility or hypothesis becomes more refined, it begins to suggest a *scenario* or *story* about what happened in the case being investigated. Event chronologies are also useful in suggesting scenarios. However, there is an important difference between a chronology, a refined possibility, and a scenario. Neither event chronologies nor possibilities outrun the evidence we have. A chronology is a simple ordering of events for which we have evidence. Refining a possibility consists of adding evidence-based details in the statement of the possibility. A scenario, however, is a mixture of what we regard as fact together with hypothetical or fanciful elements. These fanciful elements constitute the major heuristic power of constructing a scenario, as I illustrate by means of Figure 2.

Time line 1 in Figure 2 represents a scenario based on evidence items A*, B*, and C*. A* is evidence that event A occurred; B* is evidence that event B occurred, and C* is evidence that event C occurred. To tell a story involving these three items of evidence, which we refer to as *benchmarks*, we imagine a number of *gap-fillers* or *hypotheticals* that provide some explanatory connection between these three benchmarks. The gap-fillers we imagine are the events D, E, F, G, and H. We might reason, for example, that in order for event B to have occurred following A [we have evidence for each of these events], then D, E, and F might also have

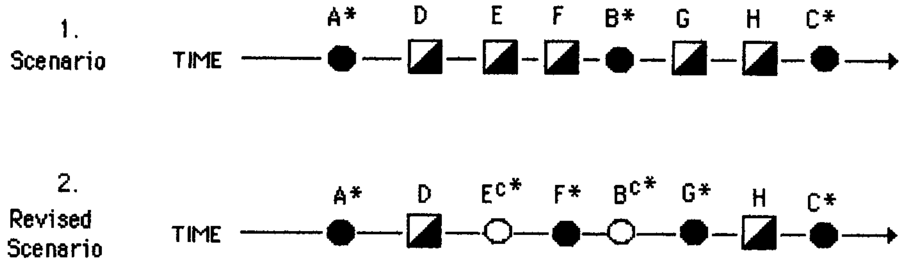

Figure 2.

happened. Similarly, we believe that hypothetical events G and H might need to be interposed between events B and C [for which we do have evidence]. One element of the heuristic merit of constructing scenarios is now exposed. We have generated five new possible lines of inquiry regarding events D, F, F, G, and H. Our belief is that there is no better heuristic device than the construction of scenarios or the telling of stories. As the continuance of this example shows, we may have to tell a different story than the one depicted in time line 1.

Time passes and we attempt to gather evidence regarding hypothetical events D, E, F, G, and H. Consider time line 2 in Figure 2. At this point, we have no evidence yet regarding events D and H [they remain hypothetical]. However, we do have evidence F^* and G^* that events F and G occurred, as we hypothesized. One trouble is that we now have evidence E^{C^*} that our hypothesized event E did not happen. In addition, we discover evidence B^{C^*}, that contradicts earlier evidence B^*. Our revised scenario now tells a different story than the one we originally told. But we must remember that having evidence about an event does not entail that this event actually occurred. Thus, we are led to examine carefully whatever credibility-related evidence we might have marshaled in our trifle base for the sources of evidence in these scenarios. This is just one example of how the user of our system will ordinarily to go from one stack to another as investigation proceeds. Credibility issues form a particularly good example of instances in which heuristic power lurks in unexpected places. Sometimes what we know about a source of evidence is at least as inferentially valuable as what this source tells us.

The stack labeled *Scenarios* in Figure 1 allows the user to construct alternative scenarios based on refined possibilities. As its cardinality [one or more] shows, the user may wish to generate one or more scenarios for each one of the possible charges or complaints being considered. Scenarios or stories are constructed for different purposes. Stories constructed by attorneys preparing for trial are intended to persuade fact finders. The stories at issue in fact investigation are those we tell for heuristic purposes in generating new hypotheses and new evidence. Our scenario stacks form an important element of the intellectual audit trails we mentioned. They provide an account of how an investigator perceived the basis for possible charges or complaints at different stages of an investigation.

Scenarios or stories eventually have themes or plots which suggest *case theories.* The case theories in legal fact investigations are grounded upon substantive legal rules. These rules prescribe the specific points or elements of a theory that must be proven, at some appropriate forensic standard of proof, for the theory to be sustained at trial. For example, suppose the scenario favored in some episode of investigation suggests the case theory: First degree murder. To sustain this theory at trial, the prosecution must show: (1) A person was killed, (2) defendant killed this person, (3) defendant intended to kill this person, and (4) this intention was premeditated. Depending upon the nature of the case, other case theories might be: Breech of contract, manslaughter, negligence, possession of a controlled substance, and so on.

Our system allows the user to marshal existing evidence under each point or element of a case theory being considered. This marshaling operation is heuristically important because it allows the user to judge how complete is the evidential coverage on all points or elements that must be proven if a case theory is to be sustained. In Figure 1 our *Case Theory* stack shows cardinality "one or more" since, during an investigation, more than one theory might be entertained. In addition, in situations in which there is a basis for more than one charge or complaint, the investigator or attorney must generate and test a case theory for each charge or complaint.

3.5. MARSHALING TIER 5: ARGUMENTS AND INFERENCE NETWORKS

At the final tier of the version of MARSHALPLAN described in Figure 1 are two marshaling operations that carry attorneys and investigators to the threshold of confrontation at trial or some other form of settlement. We first consider the use of our system by a prosecuting attorney in a criminal investigation or by an attorney representing plaintiff in a possible civil dispute. Suppose in either case that one or more case theories have been decided upon and evidence has been marshaled on the points or elements of these theories using our *Case Theories* stack. As *legal discovery* commences, prosecutors or plaintiff's attorneys begin to find out what evidence has been generated by the defense. The *Counterevidence* stack shown in Figure 1 allows prosecutors or plaintiff's attorneys to marshal the defense evidence they have been given, in accordance with legal discovery rules, that bears on each point or element of the case theories being considered by these persons. The cardinality "one or more" shows that we may have a *Counterevidence* stack for each case theory being considered. Again, there may be more than one charge or complaint being considered by a prosecutor or by plaintiff's attorney. The heuristic importance of this marshaling operation is evident. New questions, new or revised possibilities, and new evidence may be suggested by existing counterevidence.

Now suppose that an attorney has marshaled a mass of evidence he/she believes to have a bearing on the points or elements of a case theory this attorney will offer at trial. All evidence has three credentials that must be established by argument: *relevance*, *credibility*, and *probative force.* Many years ago Wigmore (1913, 1937) troubled about the difficult of establishing these credentials when there is a mass of evidence and many lines of argument to be considered. In modern terms, Wigmore's suggested solution to this problem involves the construction of *inference networks* that incorporate all available evidence together with chains of reasoning that connect the evidence with the major points or elements in a case theory. As Wigmore realized, some evidence will be directly relevant in the sense that a chain of reasoning can be constructed directly from this evidence to one of the points or elements of a case theory. But he also noted that other evidence will be only *indirectly relevant* or *ancillary* in nature. Ancillary evidence is evidence about other directly relevant evidence and the strength or weakness of links in chains of reasoning set up by directly relevant evidence. Ancillary evidence also provides a basis for probability assessments made on links in chains of reasoning when there are no relevant Statistics upon which to ground these assessments [as is usually the case in law].

Wigmore's analytic and synthetic scheme for constructing inference networks never gained acceptance among practicing attorneys. Part of the reason was that Wigmore's original methods for charting arguments are quite cumbersome; they are not "user-friendly". But Wigmore was at least sixty years ahead of his time in the study and analysis of inference networks. No one has ever challenged the logic underlying his methods. There is now a rapidly emerging technology for the probabilistic analysis of inference networks as we discuss below. Together with our colleagues, Professors Terence Anderson and William Twining (e.g., 1991), we have attempted to make Wigmore's methods for constructing inference networks more "user-friendly" (Tillers and Schum 1988). Computers can provide great assistance in constructing complex arguments from masses of evidence as we illustrate. As Wigmore realized, the construction of an inference network is an exercise in imaginative reasoning. There are no uniquely "correct" arguments. Different persons might construct different plausible chains of reasoning from the same evidence. As noted elsewhere (Kadane and Schum 1996, pp. 74–76), argument construction is a species of *abductive* reasoning.

Our evidence marshaling system incorporates the computer-based methods for argument construction we have devised. These methods are represented by the *Argument Construction* stack in Figure 1. Notice that this stack has cardinality "one or more", since there may be more than one charge or complaint and, thus, one or more case theories to be defended by argument It is important to recognize that the construction of an argument from existing evidence is also a very important heuristic device in generating new evidence and, possibly, new hypotheses. Laying out a chain of reasoning from an item of evidence to some hypothesis is also to identify sources of doubt that seem to lurk between the evidence and the hypothesis. Each source of doubt thus identified also suggests other evidence that might be gathered to reduce this doubt In fact, evidence on intermediate sources of doubt is inferentially more valuable than the evidence that set up a chain of reasoning.

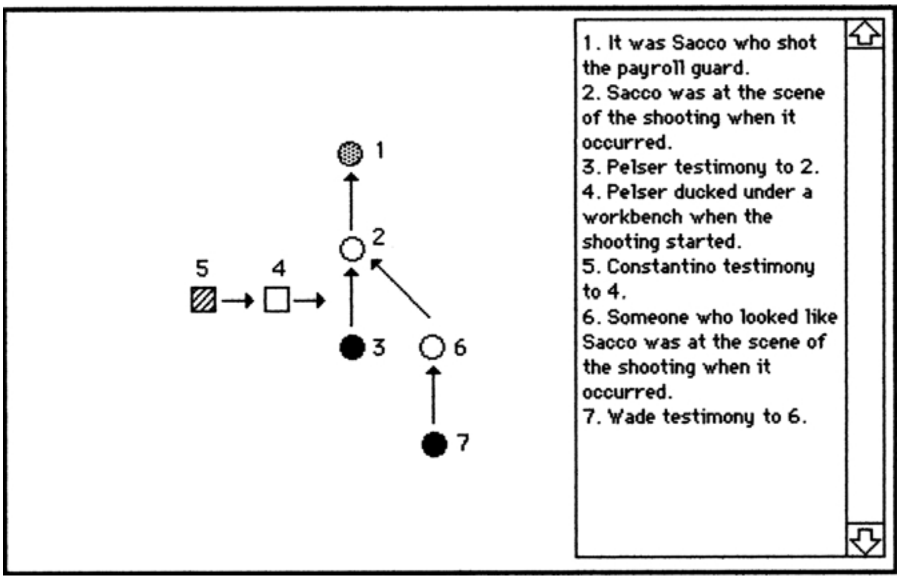

Figure 3.

The reason is that evidence bearing on links higher in a chain of reasoning will bear more directly on a hypothesis at the top of the chain.

Shown in Figure 3 is a very simple example of argument construction using our computer-based *Argument Construction* stack. This example comes from an analysis recently made of the evidence in the celebrated case of *Sacco and Vanzetti* (Kadane and Schum 1996, Appendix A). We used this *Argument Construction* stack in constructing many of the arguments based on the trial and post-trial evidence in this case. Our entire inference network for the evidence in this case consists of 28 sectors or subnetworks, each of which concerns an issue in this case. Briefly, Sacco and Vanzetti were charged with first degree murder in the robbing and shooting of a payroll guard. The two "star" prosecution witnesses against Sacco and Vanzetti were Lewis Pelser and Lewis Wade. Pelser testified that he saw Sacco at the scene of the robbery and shooting when they happened. He said he saw the crime being committed from the window of a factory building in which he worked. Wade testified that he saw "someone who looked like Sacco" at the scene of the robbery and shooting when they happened. The defense countered by producing several witnesses to impeach Pelser's credibility. One such witness was Dominic Constantino, who testified that Pelser ducked under a workbench as soon as they heard shots being fired outside.

The scrolling field to the right in Figure 3 is an example of what Wigmore called a *key list*. Key lists are collections of propositions representing evidence items, intermediate links in chains of reasoning, and major points or elements of case theories that are to be proven. This represents the *analytic* element of Wigmore's methods. Chains of reasoning, based on the items in a key list, are then constructed graphically, in terms of nodes and arcs, to show what an analyst perceives to be linkages among the propositions so listed. This graphic charting forms the *synthetic* element of Wigmore's methods. Both Pelser's and Wade's testimonies [nodes 3 and 7] are directly relevant since chains of reasoning can connect them with one of the elements in the prosecution case theory of first degree murder [node 1]: "It was Sacco who shot the payroll guard". Constantino's testimony, however, is ancillary evidence. By itself, it has no relevance on node 1, in part because it does not refer to Sacco in any way. But it acquires relevance since what Constantino said bears on Pelser's credibility and, thus, on the strength of the linkage between nodes 3 and 2. If Pelser were under the workbench when the shooting started, he cannot have observed who was at the scene of the robbery and shooting.

At this point, I can relate other features of our work on evidence marshaling with current work on KDD and data mining. The inference network shown in Figure 3 is in fact a model of how we judged Pelser's, Wade's, and Constantino's evidence to be relevant on an issue in the Sacco and Vanzetti case. But it is not a model that can be captured by any algorithm, as is the case for models generated by data mining. At the same time, however, the model in Figure 3 can be used as a basis for probabilistic analyses using any of the existing technologies for such analyses. In our work on the Sacco and Vanzetti case, for example, we employed a software system called *ERGO*™ [Noetic Systems, Inc] to perform Bayesian probabilistic analyses of inference networks we constructed using Wigmore's methods (Kadane and Schum 1996, pp. 215–239). By such means we were able to tell many different stories concerning the probative force of the evidence in this case. Systems such as *ERGO* do provide means for constructing inference networks. But they are, as yet, not able to incorporate all the subtleties in argument construction that Wigmore recognized so many years ago.

One area of current interest in data mining concerns the learning of Bayes's inference networks from experience in the form of domain knowledge and statistical data. A very good example is provided by the work of David Heckerman (1996, pp. 273–305). In many kinds of replicable situations, sources of probabilistic influence and their patterns of influence can be captured by the analytic methods Heckerman describes. Of particular interest is the capturing of *conditional nonindependencies* that may exist among sources of probabilistic influence in some process being examined. The concept of conditional nonindependence is the primary vehicle in Bayesian analyses for trapping a very wide assortment of evidential subtleties, as I have noted elsewhere (Schum 1994). So often, these subtleties are never exploited in probabilistic analyses because they are never recognized in the first place. Work such as Heckerman's offers promise that many evidential subtleties can be trapped by computers suitably equipped to learn where these subtleties may exist in large data bases. In my judgment, here is computer-assisted knowledge discovery at its finest: the trapping of evidential subtleties that may make all the difference in conclusions we must draw in so many important contexts.

But in the legal contexts we have examined, we cannot hope, at least any time soon, to have computers generate complex arguments or inference networks based on some mass of evidence to be presented at trial. The reason, mentioned several times, is that most litigation involves singular, unique, or one-of-a-kind events that are never replicable. In the construction of such networks we must rely upon the imaginative reasoning of people who have only plausibility considerations and the avoidance of non sequiturs to guide them in generating arguments in defense of the relevance, credibility, and probative force of their evidence. Wigmore's methods, I believe, offer great assistance in such difficult tasks. Methods, such as those captured in our *Argument Construction* stack, are elegant devices for marshaling our thoughts and our evidence.

Though our examples of the use of these marshaling operations were constructed from the standpoint of the prosecution or plaintiff's attorneys, these operations are equally as useful for the defense. For example, suppose legal discovery reveals certain items that are crucial to the prosecution's case. A defense attorney, charting relevance arguments from this evidence, may be led to expose sources of doubt in these arguments and can, possibly, reduce these doubts in a manner favorable to the defense. As noted earlier, all of our marshaling operations may be useful at various times by defense attorneys, particularly when their resources permit extensive investigations. Part of our charge in developing MARSHALPLAN was to design a system that could assist attorneys representing indigent clients to make the best possible use of limited investigative resources.

4. Compilations

Figure 1 shows three stacks that assist marshaling operations in various ways. All carry cardinality "exactly one", since only one is needed in any episode of fact investigation. These stacks can be easily arranged to meet the requirements of the most complex investigations. We have referred to them as *compilations*, since they allow the user to keep track of matters of interest during investigative discovery. The first compilation stack, labeled *Witnesses*, allows the user to develop lists of witnesses in a manner that may have considerable heuristic value in generating new inquiry. During an investigation an attorney frequently discovers one potential witness from the interrogation of another witness. Keeping track of when witnesses are identified, and who identified them, can be heuristically valuable. Suppose a certain witness A, who is known to witness B as having some information bearing upon a case, is later suggested by person C. Witness B never identified A in B's interrogation. In attempting to answer this question, the investigator may be led to generate new productive lines of inquiry.

Certain kinds of litigation can involve hundreds or thousands of documents; keeping track of them all in systematic ways is a chore on its own. In the stack labeled *Documents*, the user simply keeps lists of documents, not the documents themselves, arranged in ways suggested by the nature of the litigation. The final stack, labeled *Legal Rules*, is a heuristically important device that provides the user with lists of the major points or elements substantive law prescribes for various charges or complaints. Contemplating a certain case theory, the user makes reference to this stack to see just what points or elements of this case theory substantive legal rules require him/her to prove at some specified forensic standard of proof. By such means, the attorney can assess how complete or sufficient his/her evidence is at various stages of fact investigation.

5. Final words

The evidence marshaling system I have described is, in its present state, more suitable for illustrating our ideas about the importance of various evidence marshaling operations than it may be for use in actual practice. The Hypercard™ system we have used is suitable for prototyping, but is not necessarily immediately useful for application in situations that involve masses of evidence. One characteristic that evidence marshaling systems designed for application should have is a "one write" capability. What this means is that each new detail or trifle is entered into the system exactly once. A trifle or detail, once entered in our trifle base, should be easily transferable to other stacks as part of the process of examining combinations of trifles. Our present system allows only primitive means for the combining or marshaling of trifles in various ways. As we look into the future, we imagine situations in which trifles can be entered into Systems like ours by investigators who are "on the spot" or in remote locations. Such "on line" capability would enhance the speed at which fact investigation moves forward.

In spite of its present limitations, our system is useful for illustrating an important characteristic of fact investigation, case theory development, and argument structuring. In all of these activities investigators and attorneys must try to keep many things in mind at the same time, something that is very difficult for any of us to do. Systems such as ours allow these persons to keep many things at hand if not in mind. Our theory of evidence marshaling has been based on a careful study of the process of discovery in law and elsewhere and what this process seems to require as far as the marshaling or organization of thoughts and evidence. The key elements of our theory, illustrated by MARSHALPLAN, are as follows. How well we organize our existing thoughts and evidence does have a bearing on our ability to generate new hypotheses, evidence, and arguments linking them. Different stages of discovery require different marshaling operations. Every episode of fact investigation is unique and so we do not expect the came sequence of marshaling operations to be performed in every fact investigation. In law and in other contexts, sophisticated search strategies are necessary but not sufficient. Marshaling strategies that enhance the process of inquiry are equally important.

Our work relates to but is not the same as current work in knowledge discovery in data bases and in data mining. In legal contexts we face the same combinatorial explosions of data combinations and the same difficulties in extracting meaning and knowledge from combinations of data or "trifles". The uniqueness of each episode of fact investigation, and the nonreplicability of events of interest in legal fact investigations simply require strategies different from those routinely studied in KDD and data mining. In various presentations we have made of our work, we have been encouraged to believe that our system might find use in other contexts such as intelligence analysis, auditing, history, and in other contexts in which knowledge must be extracted from masses of data concerning events that are not replicable and from which no algorithmic models can be discovered. But the models our system does assist the user in generating are very useful in subsequent probabilistic analyses that may accompany the difficult process of drawing of conclusions from masses of evidence.

References

Anderson, T. and Twining, W. (1991). Analysis of Evidence: How to Do Things with Facts Based on Wigmore's Science of JudiciaI Proof. Little, Brown: Boston, MA.

Baring-Gould, W. S. (1967). The Annotated Sherlock Holmes. two volumes. Clarkson N. Potter, Inc.: NY.

Boden, M. (1991). The Creative Mind: Myths and Mechanisms. Basic Books: NY.

Carson, D. (ed.) 0000. Special Issue: United States v. Shonubi. Expert Evidence: The International Digest of Human Behavior Science and Law. Volume 4, Issue 4, 133–172.

Cohen, L. J. (1977). The Probable and the Provable. Clarendon Press: Oxford.

Cohen, L. J. (1989). An Introduction to the Philosophy of Induction and Probability. Clarendon Press: Oxford.

Cohen, M. and Nagel, E. (1934). An Introduction to Logic and Scientific Method. Harcourt, Brace and Co.: NY.

Coveny, P. and Highfield, R. (1995). Frontiers of Complexity: The Search for Order in a Chaotic Fawcett-Columbine: NY.

Eccles, J. C. (1970). Facing Reality: Philosophical Adventures by a Brain Scientist. Springer-Verlag: NY.

Eco, U. (1983). Horns, Hooves, Insteps: Some Hypotheses on Three Types of Abduction. In Eco, U. and Sebeok, T. (eds.) The Sign of Three: Dupin,. Holmes. Peirce, 198–220. Indiana University Press: Bloomington, IN.

Eco, U. and Sebeok, T. (eds.) (1983). The Sign of Three: Dupin. Holmes. Peirce. Indiana University Press: Bloomington, IN.

Eco, U. (1988). Foucault's Pendulum. Harcourt Brace Jovanovich: NY.

Einstein, A. (1954). A Testimonial from Professor Einstein. In Hadamard, J. The Psychology of Invention in the Mathematical Field, 142–143. Dover: NY.

Fayyad, U, Piatetsky-Shapiro, G., Smyth, P., and Uthurusamy, R. (eds.) (1996). Advances in Knowledge Discovery and Data Mining. AAAI Press/MIT Press: Menlo Park, CA.

Hadamard, J. (1954). The Psychology of Invention in the Mathematical Field. Dover: NY.

Heckerman, D. (1996). Bayesian Networks for Knowledge Discovery. In Fayyad, U., Piatetsky-Shapiro, G., Smyth, P., and Uthurusamy, R. (eds.) Advances in Knowledge Discovery and Data Mining, 273–305. AAAI Press/MIT Press: Menlo Park, CA.

Hintikka, J. and Hintikka, M. (1983a). Sherlock Holmes Confronts Modern Logic: Toward a Theory of Information-Seeking Through Questioning. In Eco, U., Sebeok, T. (eds.) The Sign of Three: Dupin. Holmes, Peirce, 154–169. Indiana University Press: Bloomington, IN.

Hintikka, J. (1983b). Sherlock Holmes Formalized. In Eco, U. and Sebeok, T. (eds.) The Sign of Three: Dupin, Holmes. Peirce, 170–178. Indiana University Press: Bloomington, IN.

Holland, H. I., Holyoak, K., Nisbett, R., and Thagard, P. (1989). Induction: Processes of Inference. Learning. and Discovery. MIT Press: Cambridge, MA.
Josephson, J. and Josephson, S. (1994). Abductive Inference: Computation. Philosophy. Technology. Cambridge University Press.
Kadane, J. and Schum, D. (1996). A Probabilistic Analysis of the Sacco and Vanzetti Evidence. Wiley: NY.
Kneale, W. (1952). Probability and Induction. Clarendon Press: Oxford.
Koestler, A. (1989). The Act of Creation. Arkana, Penguin Group: London.
Langley, P., Simon, H., Bradshaw, G., and Zytkow, J. (1987). Scientific Discovery: Computational Explorations of the Creative Process. MIT Press: Cambridge, MA.
Michalski, R. (1994). Inferential Theory of Learning: Developing Foundations for Multistrategy Learning. In Michalski, R. and Tecuci, G. (eds.) Machine Learning: A Multistrategy Approach, Vol. IV, Morgan Kaufmann Publishers: San Francisco, CA.
Miller, A. (1996). Insights of Genius: Imagery and Creativity in Science and Art. Springer-Verlag: NY.
Neapolitan, R. (1990). Probabilistic Reasoning in Expert Systems: Theory and Algorithms. Wiley: NY.
Oldroyd, D. (1986). The Arch of Knowledge: An Introductory Study of the History of the Philosophy and Methodology of Science. Methuen: NY.
Peirce, C. S. [1898] (1992). In Ketner, K. (ed.) Reasoning and the Logic of Things: Cambridge Conferences 1898. Harvard University Press: Cambridge, MA.
Peirce, C. S. [1901] (1955). Abduction and Induction. In Buchler, J. (ed.) Philosophical Writings of Peirce, 150–156. Dover: NY.
Peirce, C. S. [1903] (1955). Perceptual Judgments. In Buchler, J. (ed.) Philosophical Writings of Peirce, 302–309. Dover: NY.
Peng, Y. and Reggia, J. (1987). A Probabilistic Causal Model for Diagnostic Problem Solving, Parts I and II. IEEE Transactions on Systems. Man. and Cybernetics, vol. SME-17.
Penrose, R. (1991). The Emperor's New Mind: Concerning Computers. Minds, and the Laws of Physics. Penguin Books: NY.
Platinum Technology. Paradigm Plus (1996). Methods Manual. Release 3.5.
Poincaré, H. [1908] (1956). Mathematical Creation. In Newman, J. (ed.), The World of Semantics. Vol. IV. Simon and Schuster: NY.
Popper, K. (1968). The Logic of Scientific Discovery. Harper Torchbooks: NY.
Price Waterhouse Technology Forecast: 1966. (1996) Price Waterhouse World Technology Centre: Menlo Park, CA.
Reichenbach, H. (1968). The Rise of Scientific Philosophy. University of California Press: Berkeley, CA.
Schum, D. (1989). Knowledge, Credibility, and Probability. Journal of Behavioral Decision 2: 39–62.
Schum. D. (1992). Hearsay from a Layperson. Cardozo Law Review 14(1): 1–77.
Schum, D. (1994). Evidential Foundations of Probabilistic Reasoning. Wiley: NY.
Schum, D. and Tillers, P. (1991). Marshaling Evidence for Adversary Litigation. Cardozo Law Review 13(2–3): 657–704.
Simon, H. (1973). Does Scientific Discovery have a Logic? Philosophy of Science 40: 471–480.
Thagard, P. (1993). Computational Philosophy of Science. MIT Press: Cambridge, MA.
Tillers, P. and Schum, D. (1988). Charting New Territory in Judicial Proof: Beyond Wigmore. Cardozo Law Review 9(3): 907–966,
Tillers, P. and Schum, D. (1991). A Theory of Preliminary Fact Investigation. University of California Davis Law Review 24(4), 931–1012.
Tursman, R. (1987). Peirce's Theory of Scientific Discovery: A System of Logic Conceived as Semiotic. Indiana University Press: Bloomington, IN.
Wigmore, J. H. (1913). The Problem of Proof. Illinois Law Review 8(2), 77–103,

Wigmore, J. H. (1937). The Science of Judicial Proof: As Given by Logic. Psychology, and General Experience and Illustrated in Judicial Trials. 3rd ed. Little, Brown: Boston, MA.
Wolpert, L. (1992). The Unnatural Nature of Science. Harvard University Press: Cambridge, MA.