

# Empirically Quantifying Evidence Assessment in Legal Decisions

Vern R. Walker

Professor of Law  
Director, Research Laboratory for Law, Logic and Technology  
Hofstra University School of Law  
121 Hofstra University  
Hempstead, New York 11549, USA  
{Vern.R.Walker@Hofstra.edu}

**Abstract.** This paper reports the baseline framework and results for the Vaccine/Injury Project at Hofstra Law School’s Research Laboratory for Law, Logic and Technology (LLT Lab). The Project models the fact-finding reasoning in compensation decisions about whether a vaccination caused a particular adverse medical condition. Results reported involve (1) protocols for modeling the reasoning found in natural-language decisions (e.g., modeling legal rules, evidentiary assertions, findings and evidentiary structure) and (2) patterns of reasoning found in the models (e.g., the influence of rule content on evidence structure, the role of policy in fact finding, the development of “soft rules” of inference, and possible quantification metrics).

**Keywords:** legal reasoning, legal rule, logic model, plausibility, vaccine compensation, artificial intelligence

## 1 Introduction

Since Wigmore’s early work on evidence charting, a core problem has been developing a methodology for modeling and grading the “force of evidence” or “evidential support” for findings of fact in legal decisions [1, 2]. There are theories of formal structures for default or defeasible inference [3, 4, 5], as a sub-area of nonmonotonic logic [6], as well as informal logic theory in the form of argumentation schemes [7, 8]. Moreover, artificial intelligence research applied to law employs default logic [9, 10, 11] and argumentation schemes [7, 12, 13], and has been investigating structures useful in fact finding [14, 15, 16, 17]. The nature of law, however, suggests that such theoretical work should be complemented by analyses of actual factfinder reasoning about complex factual issues in documented legal decisions, preferably ones involving both scientific and non-scientific evidence.

The fact-finding processes found in law are distinctive decision-making processes because they require a balancing of the epistemic objective against non-epistemic objectives [18, 19]. The epistemic objective is to produce findings of fact that are as accurate as possible, as well as warranted by the evidence available to the factfinder. Non-epistemic objectives can be both generic (such as procedural fairness to parties

and administrative efficiency) and specific (such as protecting public health from unsafe food). Studying the actual balance struck in different types of legal proceeding requires a methodology for accurately modeling the fact-finding reasoning found in natural-language legal decisions – a methodology that is grounded in theory, empirically refined and tested, and generalizable to different areas of decision making. This paper reports on current research that is developing techniques for reliably and validly analyzing the logical structure of the reasoning of a factfinder, while also developing a database of logic models of complex legal decisions.

The techniques for modeling fact-finding reasoning should be flexible enough to apply to any decision written in a natural-language document, should produce models (data) that are standard in form, and should be empirically tested for reliability and validity. This suggests three performance criteria for such models. The first is descriptive accuracy – namely, that a model for a decision must accurately capture the logical structure of the reasoning as reported by the decisionmaker. The second is normative completeness – that is, that the model must contain sufficient information about the reported reasoning, so that a normative critique of the model is also a valid critique of the reported reasoning. The third criterion is fitness for automation – it must contain, if at all possible, structures and functionalities that can be automated.

## 2 The Vaccine/Injury Project of the LLT Research Laboratory

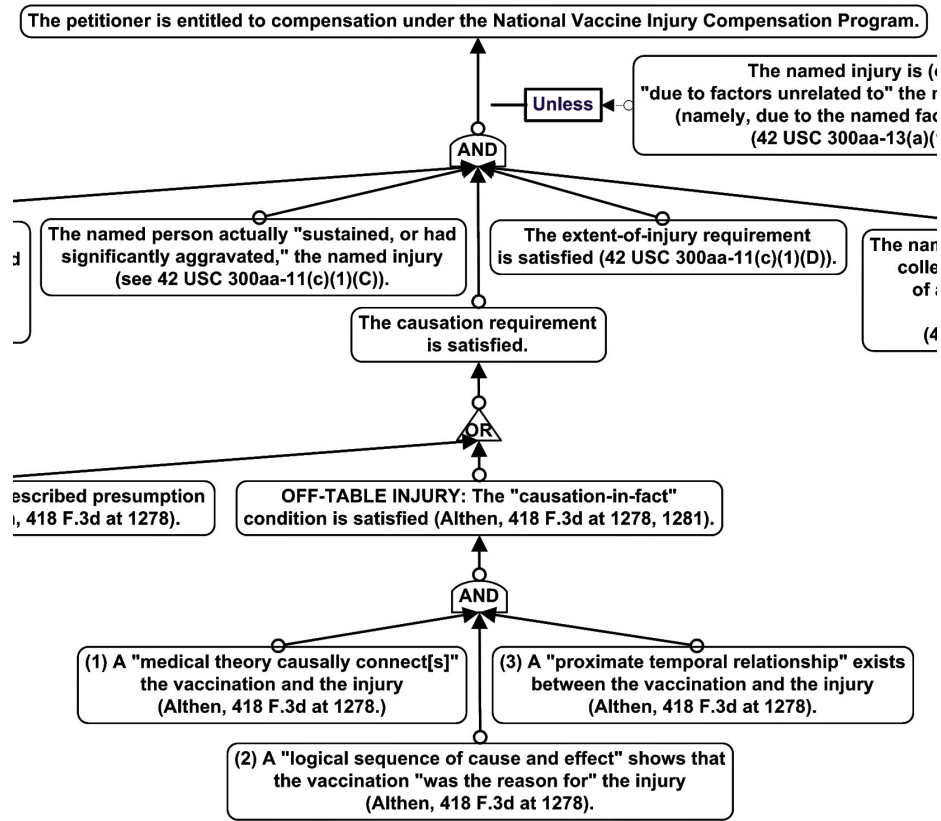
This paper reports on current empirical research at Hofstra Law School’s Research Laboratory for Law, Logic and Technology (LLT Lab). The Lab’s mission is “to conduct empirical research on substantive areas of law using a logic-based analytic framework and state-of-the-art technology, thereby creating knowledge, skills and tools useful in both legal practice and legal education.” Primary goals of the LLT Lab are to create a legal research laboratory modeled on research laboratories in the sciences; to use teams of faculty and students to develop and disseminate practical tools for legal research, education and practice; to develop methodologies that enable the upward scaling of research projects by including additional researchers and collaborating legal research laboratories; and to use technology to produce useful knowledge and tools for society.

In particular, this paper reports some of the baseline framework and results of the Vaccine/Injury Project (V/IP), which is investigating the claims process in the United States for compensation due to injuries caused by vaccinations. The Vaccine Injury Compensation Program (VICP) is a hybrid administrative-judicial system in the United States for providing compensation to persons who have sustained vaccine-related injuries [20]. Compensation awards are paid out of the Vaccine Injury Compensation Trust Fund, which is funded by an excise tax on each dose of covered vaccine. Petitioners claiming compensation file claims in the United States Court of Federal Claims. Petitions contested by the Secretary of the Department of Health and Human Services (HHS) are decided by one of eight special masters in the Office of Special Masters, which Congress established for this purpose within the Court of Federal Claims. Contested VICP cases often involve complex issues of fact, which require taking into account medical, scientific, and other expert evidence along with non-expert evidence. Moreover, the special masters carefully document their evidence

assessment and fact finding in their decisions, which must include “findings of fact and conclusions of law.” Under certain conditions, the Court of Federal Claims has jurisdiction to review a special master’s decision, and a judgment of that Court can be appealed to the United States Court of Appeals for the Federal Circuit. However, the findings of fact made by special masters may be set aside only if they are “arbitrary and capricious” – a very deferential standard of review [20]. Such factfinders have wide discretion in assessing the probative value of the evidence and in drawing conclusions about the facts.

The study sample of decisions selected for the project consists of decisions filed by special masters between 29 July 2005 and 18 June 2009. The sample start date is the date of the Federal Circuit decision in the lead case of *Althen v. Secretary of HHS*, 418 F.3d 1274 (July 29, 2005), which established a 3-prong test of causation-in-fact. The sample end date is the date of the Federal Circuit decision in *Andreu v. Secretary of HHS*, 569 F.3d 1367 (Fed. Cir. 2009). The study sample consists of every decision filed by a special master during this time frame, and in which the special master applies the 3-prong test of causation-in-fact. This test states conditions for proving (in certain complex cases) whether vaccinations have caused injuries. This study sample was selected for its societal importance, for its complex yet stable legal rules concerning proof of causation, and for the size and representativeness of the sample. When completed, the study will contain at least 100 decisions.

The LLT Lab utilizes the Default-Logic Paradigm as the formal methodology for its modeling [18, 21, 22]. This requires modeling the logical structure of the legal rules on causation applicable during the study sample. The Default-Logic Paradigm models systems of rules as inverted “rule trees.” Figure 1 shows part of the rule tree for the VICP, including the 3 sub-issues that the petitioner must prove under the *Althen* rule for causation: (1) that a “medical theory causally connects” the vaccination and the injury; (2) that a “logical sequence of cause and effect” shows that the vaccination “was the reason for” the injury; and (3) that a “proximate temporal relationship” exists between the vaccination and the injury. (*Althen*, 418 F.3d at 1278.) The top node of the tree is the ultimate issue to be proved by the petitioner. Each level of each branch extending downward from the top node states the logical conditions for proving the immediately higher proposition. Inferences proceed upward. Figure 1 also illustrates three logical connectives: “AND” (the conclusion is true if but only if all connected conditions are true); “OR” (the conclusion is true if but only if at least one of the connected conditions is true); and “UNLESS” (if the defeating condition is true, then the conclusion is false, even if the main *prima facie* conditions are true). A fourth logical connective used to model legal rules is “RULE FACTORS,” which is used when the rule requires the factfinder to take the listed factors into account when making a finding, but does not specify an algorithm for assigning a truth-value to the conclusion as a function of the truth-values of those factors. The shape of a rule tree tends to be triangular, with the single ultimate issue as the apex at the top, which is dependent for its truth or falsehood upon combinations of factual issues that terminate the branches along the triangle’s base at the bottom.



**Figure 1. Part of the VICP Rule Tree, Containing the Three *Althen* Sub-Issues for Causation-in-fact, and Illustrating Three Logical Connectives**

Every proposition in a rule tree has one of three values: “true” / “undecided” / “false”. When a legal proceeding begins, all propositions in the rule tree are “undecided.” In any particular vaccine-compensation case, the petitioner and the Secretary of HHS produce evidence and try to prove or disprove the applicable issues of fact. The special master’s role as factfinder is deciding which evidence is relevant to which issues of fact, evaluating the plausibility of the evidence, organizing the evidence and making reasonable inferences, and making findings of fact. Often this requires deciding between two conflicting expert opinions or integrating expert and non-expert evidence into a single line of reasoning. When the factfinder makes findings of fact about the “terminal propositions” at the ends of the branches, the tree’s logical connectives propagate those values upward, determining the new values of propositions throughout the tree, including the truth-value of the ultimate issue at the top (in vaccine cases, either awarding compensation or not).

In modeling the reported reasoning in any particular decision, Lab researchers first select the rule tree containing the legal rules governing the case. They then begin to model the evidence assessment, by extracting from the decision all assertions that the

special master reported as relevant to the three conditions of causation. Such assertions may be statements made by testifying witnesses, statements contained in documents, or statements of the special master (e.g., describing an evidentiary exhibit or a witness's demeanor). The special master's findings of fact are also assertions. In deciding a case, the special master explicitly or implicitly organizes these assertions and assigns them degrees of plausibility. The Vaccine/Injury Project models this assignment using an ordinal, seven-valued scale, with the values being: "highly plausible" / "very plausible" / "slightly plausible" / "undecided" / "slightly implausible" / "very implausible" / "highly implausible". In vaccine-compensation cases, where findings are made by a preponderance of the evidence, an ordinal scale with seven values provides sufficient evaluative precision [18].

The next step in modeling the evidence assessment is to identify the special master's findings of fact on causation and to attach them to the appropriate propositions in the rule tree. Then researchers attach to those findings the special master's reasoning, organizing the extracted evidentiary assertions into logical structures using "plausibility connectives." Plausibility connectives are similar to logical connectives, but operate on seven plausibility-values. The plausibility counterpart to AND is "MIN," which assigns to the conclusion the *lowest* plausibility-value possessed by any of its conditions (premises). The "MAX" connective (the counterpart to OR) assigns to the conclusion the *highest* plausibility-value possessed by any of its conditions (premises). The plausibility connective "REBUT" is the counterpart to UNLESS. If the rebutting assertion is plausible to any degree, then the REBUT connective assigns to the conclusion a degree of *implausibility* inverse to the degree of plausibility of the rebutting (defeating) assertion. For example, if the rebutting defeater is "highly plausible," then the conclusion is "highly implausible"; but if the defeater is only "slightly plausible," then the conclusion is only "slightly implausible." The project also uses the connective "EVIDENCE FACTORS" to model reasoning where the special master indicates that she or he took those assertions into account, but where the decision does not provide information about how the special master organized the assertions to produce a logical chain of reasoning. Finally, researchers assign plausibility-values to the terminal assertions in the model, as well as to sets of evidence factors, so that the plausibility-values and truth-values throughout the model reflect the evaluation of the factfinder.

In producing logic models of fact-finding reasoning, the LLT Lab uses the Legal Apprentice software from Apprentice Systems, Inc. The software creates XML-formatted files of the logic models, which is a standard format used in Internet-based programs. The project also makes such models available in HTML format. Figure 2 is a screen shot of the Legal Apprentice model for the vaccine decision *Casey v. Secretary of Health and Human Services*, Case No. 97-612V (December 12, 2005). On a color computer display or a page printed in color, the round icon before each assertion has a color that indicates the plausibility-value of the assertion on that line. Figure 2 illustrates a finding of the special master that "there is an adequate medical theory of causation," supported by two alternative lines of reasoning (connected by the plausibility connective MAX) based on two causal pathways (direct viral infection and immune-mediated inflammatory response).

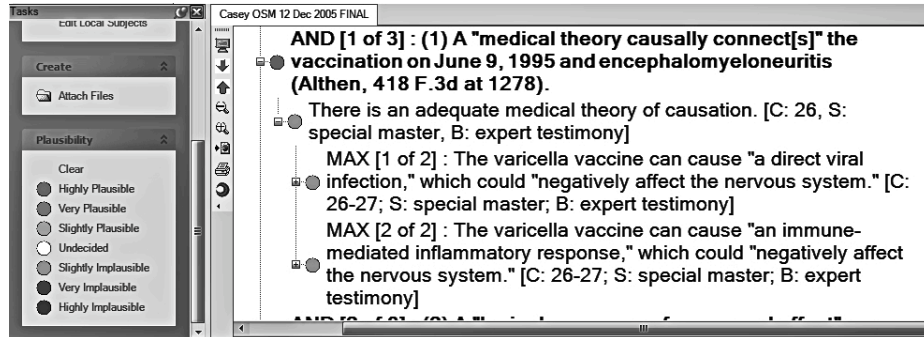


Figure 2. Illustration of a Portion of the Logic Model for the *Casey* Decision in the Legal Apprentice Software

### 3 Toward Quantifying Modeling Protocols

A major objective of the Vaccine/Injury Project is to develop and test protocols for reliably and validly generating logic models for legal documents. “Reliability” in this context means the degree of consistency in models when different researchers model the same decision, and “validity” means the degree to which a model accurately captures the reasoning reported by the factfinder [23]. For determining the *validity* of the logic modeling, the project relies primarily upon the Principal Investigator’s expertise and experience in logical theory, in the use of scientific evidence in law, and in legal fact-finding processes. Because the evidence and reasoning in each decision is unique, the project cannot employ other methods of establishing validity, such as criterion, content or construct validity [24].

For developing and testing *reliable* protocols for logic models, the project assigns a pair of researchers to each decision in the study sample. One researcher (“the modeler”) reads the decision; extracts the relevant evidentiary assertions; tags those assertions as to citation, source, and other factors; and connects those assertions into a model of the factfinder’s reasoning. The second researcher (“the reviewer”) independently checks this model against the original decision. The two researchers then confer on any differences of opinion, document the nature of those differences, and reach a consensus. Finally, the Principal Investigator and the two researchers discuss and document any issues and decisions, and then decide on the final model.

This process provides experience and documentation that is used to refine protocols for modeling. The project will employ standard designs and statistics for testing the reliability of the resulting modeling protocols. It will design a reliability study from a wide variety of English passages drawn from actual decisions, will test the consistency of modeling across multiple modelers, and will quantify and report the results. If protocols can be developed that will produce acceptably reliable and valid logic models for legal decisions written by a variety of authors in normal English, then such protocols will be useful for training new researchers, for providing quality assurance in modeling, and for guiding the automation of the modeling

process. The following subsections discuss some of the baseline results from this work on protocol development.

### 3.1 Identifying and Modeling Legal Rules

Legal rules are relatively easy to define in terms of logical structure, but less easy to locate in legal decisions. Legal rules are conditional propositions that have been authoritatively adopted as universally applicable [17, 18, 25]. A conditional proposition has the logical form “if  $p$ , then  $q$ ,” where  $p$  and  $q$  stand for two constituent propositions, and  $p$  can itself contain constituent propositions connected by a logical connective. A rule is a statement that finding proposition  $p$  (the condition) to be true warrants finding proposition  $q$  (the conclusion) to be true also. As described in Section 2, the Default-Logic Paradigm models sets of legal rules as inverted rule trees consisting of three-valued propositions as nodes and three-valued logical connectives.

The definition of a legal rule provides several criteria for locating rules in a legal decision: (i) they are conditional propositions; (ii) they have been authoritatively adopted; and (iii) they are intended to be universally applicable. In addition, the fact that rules connect together in sets is an important attribute. These four characteristics will be discussed in their order of usefulness in searching documents for legal rules.

First, there is a convention to cite the relevant legal authority when stating a legal rule. In U.S. legal decisions, for example, the legal authority is often cited immediately after the rule, often using a standard format. For example:

**Example 3.1a:** To make the requisite showing, petitioner must offer “proof of a logical sequence of cause and effect showing that the vaccination was the reason for the injury.” Shyface v. Sec’y of HHS, 165 F.3d 1344, 1353 (Fed. Cir. 1999).

(Such examples in this paper are drawn from the *Casey* decision, cited above.)

Second, because legal rules are intended to be universally applicable, they are generally stated using only indefinite logical subjects, without any definite subjects. Indefinite subjects refer to objects or events only as a class or by their general characteristics, and are usually expressed in English by common nouns (“cerebellitis”), indefinite descriptions (“an autoimmune disease”), and other grammatical phrases that name groups or classes of individuals. By contrast, definite subjects denote specific individuals, as by proper names (“Ms. Casey”) or by definite descriptions (“the special master”). Pronouns can denote either indefinite or definite subjects, depending upon the context. In Example 3.1a, the noun phrase “a logical sequence of cause and effect” is indefinite. On occasion, however, the writer of a decision in English will efficiently state a rule by substituting definite subjects for the normal indefinite subjects, which poses obvious problems for eventual automation.

Third, while it must always be possible to reformulate a statement of a legal rule in English into its logical conditional form, this sometimes requires a great deal of interpretation. Occasionally the decision’s drafter will use the words “if” and “then” in an informative way to signal the condition(s) and conclusion of a rule, but this is far from always the case, as the following example illustrates:

**Example 3.1b:** Without more, “evidence showing an absence of other causes does not meet petitioners’ affirmative duty to show actual or legal causation.” Grant, 956 F.2d at 1149.

In this example, the rule to be abstracted is that *if* a petitioner in a vaccine compensation case only produces “evidence showing an absence of other causes,” *then* that petitioner has necessarily failed to prove causation, as a matter of law.

Fourth, the fact that rules occur in integrated sets or systems is an important characteristic in identifying rules. For example, some rules provide definitions for legal terms that occur in other rules. In modeling such a rule in the Default-Logic Paradigm, the defining conditions (the *definiens*) occur in the immediately lower level of the rule tree, connected to the conclusion (the *definiendum*) by a logical connective (in the case of definitions, usually AND). As a search strategy for identifying sentences that contain rules, therefore, the words that occur in the terminal propositions of the rule tree provide important search words for identifying any additions to those rules that might occur in the decision.

### 3.2 Identifying and Modeling Evidentiary Assertions

As with legal rules, the basic building blocks of evidence assessment are propositions (statements capable of being either true or false), but the Default-Logic Paradigm models them using seven ordered plausibility-values instead of three truth-values, and calls them “evidentiary assertions,” or simply “assertions.” Examples of evidentiary assertions are statements made by testifying witnesses, statements contained in documents that are admitted into evidence, and statements of the factfinder summarizing or characterizing the evidence, as well as the findings of the factfinder. Unlike legal rules, evidentiary assertions do not have a single logical form that can be used for identifying them, and they can contain references to either definite or indefinite subjects. Evidentiary assertions can also contain legal terms (particularly within findings of fact), corresponding to the concept of “mixed issues of law and fact.” A working hypothesis at the LLT Lab is that any sentence or clause in a decision that does not state a legal rule, a legal policy, or a legal principle is modeled as an evidentiary assertion.

One inferentially important type of evidentiary assertion is a “generalization” [1, 3, 8, 18]. Examples are: “most witnesses testifying under oath tell the truth,” “one-third of Americans are overweight,” and “60% of the test group in the study developed the disease” – which have the following logical forms (respectively): “most As are Bs,” “X/Y of As are Bs,” and “X% of the members of group A are (also) members of group B.” A generalization typically includes an asserted degree of logical quantification over the reference class, as well as an explicit modal “hedge” qualifying the entire assertion – such as expressions of frequency (“often”), typicality (“normally”), temporal limitation (“in the past”), or degree of confidence (“perhaps”) [18]. A lexicon of such words might provide useful search tools for generalizations, and help to distinguish generalizations from legal rules (which apply universally).

An important grammatical aspect of English and other natural languages is the embedding of propositions within propositions. Embedded propositions are constituents of other propositions, for example:

**Example 3.2:** However, Dr. Tornatore stated that it is reasonable to assume that the virus in the vaccine still can multiply.

In this example, the clause “the virus in the vaccine still can multiply” is the substantive assertion of the evidentiary reasoning, which is embedded in the clause “it is reasonable to assume that ...” (stating a modal hedge), which is embedded in turn in the main sentence “however, Dr. Tornatore stated that ...” (stating the source of the assertion). The model for the *Casey* decision includes the embedded substantive statement as the assertion, annotated or tagged with the remaining information: “The virus in the vaccine still can multiply. [Source: Dr. Tornatore; Modality: “it is reasonable to assume”].” In modeling evidentiary reasoning, it is important to determine whether the primary proposition or the embedded proposition provides the content for the reasoning. In some contexts, the important evidentiary information might be the fact that the witness made the statement, rather than the statement itself.

### 3.3 Modeling Findings and the Core Evidentiary Structure

A critical subset of the assertions in a decision consists of the factfinder’s findings of fact. The factual issues are identified by the terminal propositions at the ends of all the branches of the rule tree. The modeling of the evidence assessment in a particular case begins by identifying and modeling all of the findings of fact on those issues.

Identification of a finding of fact is greatly assisted if the decision author uses such ritualistic words as “finds” or “finding,” or related words with similar meanings, such as “determines” or “concludes.” However, this is a matter of the author’s style. Even if the author’s general practice is to signal findings using such words, this is no guarantee that searching for such words will identify all of the findings in a decision.

Authors of legal decisions are often careful to formulate their findings of fact using the legal terminology of the rule condition to which the finding relates. A useful indication of a finding is that it contains both a legal phrase or concept from a terminal proposition in the rule tree, as well as definite subjects referring to specific persons or events in the particular case (which a rule statement itself would seldom contain). As with other modeling issues, the variety of author styles and the variety of English grammatical forms provide a challenge to automation, but automation can still assist the human modeler by utilizing rule terms to provide search hits.

After identifying a finding of fact in a decision that corresponds to a particular rule condition, the question is what model structure to use to attach the finding to the rule tree. The paramount goal is achieving descriptive accuracy with respect to the decision. With the majority of contested issues of fact, there is a standard approach to the modeling. Typically, the party that has the burden of proving the issue of fact (or a witness on behalf of that party) asserts that the condition is true (satisfied), while the opposing party (or a witness on behalf of that party) asserts that it is not true, and the factfinder finds in favor of one of these two sides. Thus, for contested issues of fact, the standard model has the core structure shown in Example 3.3, and the finding of fact is either the assertion of the proponent or of the opponent.

**Example 3.3:** {Contested rule condition}  
                   {Assertion of the proponent}  
**REBUT** {Assertion of the opponent}

There are variations on this core structure, however, particularly when the factfinder decides against the proponent of the issue. For example, in some cases it is more accurate to attach the proponent's allegation to the rule condition, then model the proponent's supporting reasoning (e.g., a list of alternative MAX arguments), and attach the opponent's and factfinder's REBUT defeaters to one or more of the proponent's supporting reasons. If the proponent has the burden of proof, then a finding against the proponent is warranted if he fails to supply convincing evidence.

### 3.4 Modeling Relevance and Selecting Plausibility Connectives

Once findings are identified and appropriately attached to the branches of the rule tree, the next task is to model the supporting reasoning for those findings, which includes modeling the arguments urged by the parties. Whenever the supporting reasoning involves two or more conditions on the same level, this requires selecting a plausibility connective for combining them. This section discusses a strategy for selecting plausibility connectives for a model.

Many decision authors state as supporting reasons the factors that they consider relevant and took into account in reaching the finding, but they do not report how they combined those factors to reach the conclusion. In such situations, valid modeling requires use of the connective EVIDENCE FACTORS, which allows the modeler to attach the stated reasons and evaluate them with respect to plausibility, but contains no algorithm for assigning a plausibility to the conclusion as a function of the plausibility of the factors. To use a stronger connective in the model may be unwarranted given the reported reasoning. However, listing relevant factors records in the model important information for normative critique and future development, and allows more reasoning to be added in support of individual factors.

Whenever warranted by the decision, however, the plausibility connectives MIN and MAX should be used in lieu of EVIDENCE FACTORS because they contain inferential information that is otherwise lost. The MAX connective occurs less frequently in our vaccine models than MIN, because MAX is reserved for connecting two or more alternative and independently sufficient lines of argument leading to the same conclusion. Syllogistic and other sets of necessary premises support their conclusions by assertions connected together by MIN. The connective REBUT is used only when the assertion defeats the conclusion. In general, therefore, EVIDENCE FACTORS is the weakest and always available connective within reasoning, and is replaced by MIN whenever an argument has necessary elements, with MAX and REBUT being used only when specifically warranted.

It is an open hypothesis in the LLT Lab's research whether additional plausibility connectives should be defined and utilized. For example, while the connective UNDERCUT is well defined [10, 12, 18], we have not yet encountered the need to use it in our modeling.

## 4 Toward Quantifying Reasoning Patterns

Once descriptively accurate models have been constructed, a major objective is to abstract from the logic models those patterns of reasoning that do or might re-occur, and to develop “plausibility schemas” based on those patterns [18]. In logic, a “schema” is a formal linguistic pattern containing variables, such that appropriate substitutions for the variables create instances of the pattern. “Plausibility” schemas are general patterns of evidentiary reasoning, consisting of evidentiary assertions and plausibility connectives. When a plausibility schema is instantiated with substitutions for the particular case, it warrants a default inference that the conclusion is plausible, provided the assertions that function as premises are plausible. Plausibility schemas state the conditions under which the residual uncertainty (potential for error) in making the inference is acceptable. Plausibility schemas can furnish the basis for guiding and critiquing the reasoning of factfinders in future cases, by explaining how a *reasonable* factfinder might assess the available evidence and make findings of fact. They also increase the likelihood of automating various sub-tasks of fact finding. Many of the schemas and the insights into policy-based reasoning should also be transferable to other areas of law, as well as to decision making in non-legal areas.

### 4.1 The Influence of Rule Content on Evidence Structure

Just as rule conditions inform the content of the issues to be decided and the findings to be made, the content of a rule condition sets constraints of relevance on all assertions properly attached to that node or any node beneath it. Propositions as tree nodes define “branches of relevance” centered on a single issue of fact. Understanding the meaning of the terminal proposition in a rule tree guides determinations of relevance for attaching evidence to it, and monitoring what evidence the factfinders attach to what terminal propositions informs us as to their understanding of the meaning of the rules.

In the LLT Lab’s Vaccine/Injury Project, for example, we focus our modeling attention on the proof of causation-in-fact. Of the three causation prongs or conditions in the petitioner’s *prima facie* case that the *Althen* case established (see Section 2 above), the case law and the understanding of the special masters is that the first condition to be proved (namely, that a “medical theory causally connects” the vaccination and the injury) concerns “general causation” – i.e., whether the type of vaccine can in fact cause the type of injury alleged. The second and third conditions (that a “logical sequence of cause and effect” shows that the vaccination “was the reason for” the injury, and that a “proximate temporal relationship” exists between the vaccination and the injury) establish “specific causation” – i.e., whether the vaccination in the particular case actually caused the particular injury. The first prong is about the possibility of causation in general, while the second and third prongs are about causation in the specific case.

This understanding of the three prongs provides important information for sorting assertions under them by relevance, and for anticipating types of patterns we expect to see under each. For example, assertions relevant to the first prong (general causation) tend to contain only indefinite subjects, while assertions relevant to the second and

third prongs (specific causation) tend to refer to definite individuals and events in the specific case. While these are not universal rules, they are useful guidelines.

Moreover, lines of reasoning supporting the first prong (general causation) tend to articulate causal models that rest for their validity only upon published studies and accepted scientific theories. The supporting patterns of reasoning tend to be those encountered in the scientific and medical literature and utilized by scientists themselves. By contrast, arguments under the second prong (specific causation) are about what should count as sufficiently probative evidence that the causal model developed in the first prong adequately describes what occurred in the specific case. This is more clearly outside the realm of science, and should involve compensation policies under the Vaccine Act. This example illustrates how the content of rule conditions informs the patterns of reasoning that are relevant to proving them, and how the acceptance of those patterns by factfinders tends to inform the meaning of the rules in later cases. The nature of different branches in the rule tree also informs how we search for relevant reasoning.

#### 4.2 The Role of Policy in Fact Finding

The example just above about the three branches or prongs of the *Althen* test of causation also suggests an hypothesis about balancing the epistemic objective (accurate and warranted findings) against non-epistemic objectives (e.g., encouraging vaccine production, administrative efficiency). The hypothesis is that the special masters, as knowledgeable and repeat factfinders, would explain their fact finding under the first prong of *Althen* (general causation) using the rationales generally accepted by scientists, because there might be a tendency to view general causation as entirely scientific. By contrast, under the second prong (specific causation), we might expect the special masters to use their discretion more to implement policy, because the conclusion at issue is not provable scientifically in the same way.

In fact in the vaccine decisions we find an interesting interplay between fact finding and policy implementation. As the *Casey* decision explained:

In essence, the special master is looking for a reputable medical explanation of a logical sequence of cause and effect (Grant, 956 F.2d at 1148), and medical probability rather than certainty (Knudsen, 35 F.3d at 548-49). As the Federal Circuit explained in Knudsen, medical probability means biologic credibility or plausibility: “Causation in fact under the Vaccine Act is thus based on the circumstances of the particular case, having no hard and fast per se scientific or medical rules.” 35 F.3d at 547.

*Casey*, at 9. Discovering how special masters translate such policies into fact finding requires empirical investigation.

When the data display fact-finding patterns, the Vaccine/Injury Project tries to develop plausibility schemas useful in future cases. For example, a preliminary plausibility schema for fact finding under *Althen* prong 1, based on reasoning in *Casey*, is shown in Example 4.2 (variables for substituting subjects are enclosed in {} brackets):

**Example 4.2:**

{The named vaccine} can cause {the named medical condition}.

**MIN [1 of 2]** {The named vaccine} contains a live but attenuated virus.

**MIN [2 of 2]** A wild virus of the same type as the virus contained in {the named vaccine} can cause {the named medical condition}.

**REBUT** There is sufficiently probative defeating evidence.

**MAX [1 of 2]** The virus in {the named vaccine} cannot reproduce once inside the human body.

**MAX [2 of 2]** The theoretical possibility of causation is disproved.

When subjects are substituted for variables in a particular case and the resulting *prima facie* premises are plausible, this schema produces a warranted inference and plausible conclusion. For example, in *Casey*, where the special master found for the petitioner, the named vaccine was the varicella vaccine and the named medical condition was a direct viral infection; the petitioner's expert established the two MIN assertions (the *prima facie* case), while the government's expert did not dispute the theoretical possibility of causation, but merely contended that such a reaction would occur very rarely. This schema and reasoning reflects a pro-petitioner policy, and does not impose on the petitioner a scientifically rigorous standard of proof. It is an empirical question how such an argument *in fact fares* in other cases before other special masters, and a normative question how it *should fare*, given the policies of the Vaccine Act.

### 4.3 The Development of "Soft Rules" of Inference

A major hypothesis being investigated in the LLT Lab is the development and function of "soft rules" of inference – that is, generalizable patterns of reasoning that have become "safe havens" of inference because they have been affirmed by legal authority to be reasonable [18]. Soft rules occur when a supervisory authority (in the Vaccine/Injury Project, the judges of the Court of Federal Claims or the Federal Circuit Court of Appeals) decides that a particular finding is a reasonable inference from the particular evidence found in the particular case. The hypothesis is that in an area of complex cases (such as the vaccine-compensation cases), with a small number of repeat factfinders (the special masters), and documentation of the supervisory decision once it occurs (the court judgments), at least some patterns determined by authority to be reasonable would become "safe havens" for factfinders who do not wish to be reversed and who have reason to be efficient. The hypothesis is that we will see similar inferences re-occur in later decisions, patterned on the soft rules.

### 4.4 Developing Quantification Metrics for Logical Structure

One important objective of the LLT Lab is to generate and make available to the public databases of legal decisions and corresponding logic models, so that other

institutions and sciences can perform they own analyses. For example, Apprentice Systems, Inc. has begun development of quantification metrics based on the Lab's Vaccine/Injury Project.

Some quantification metrics address the complexity of rule trees. One example under development is a Formal Complexity Index for rule trees, heavily influenced by the number of terminal propositions that must be established in order to create a winning argument, together with the number of available defeaters. Another example is a Variability Index influenced by the number of alternative ways a successful case can be proved. Such indices of complexity would allow quantitative comparison of different rule systems, or modeling the evolution of the same rule system over time.

Quantification metrics are also being developed for the logical structure of evidence assessment. For example, failure rates and variability can flag critical nodes where cases are often won or lost, and suggest investigations into the patterns that lead to success or failure. Another example is performing quantitative semantic analyses on the content of the models, and comparing those results with semantic analyses for the natural-language decisions. Such analyses could also indicate combinations of decisions where new patterns of reasoning are emerging.

Such quantitative tools will grow in importance as the number of decisions in a database grows, and as comparative analyses with new databases become possible.

## 5 Conclusion

This paper reports considerable progress toward developing empirical methods for modeling the logic of natural-language legal decisions, as well as the eventual quantification and automation of that modeling. In addition, the LLT Lab's protocols for modeling, databases of modeled legal decisions, and libraries of reasoning patterns should provide useful resources for formal and informal logic theory, for natural-language research in linguistics (especially semantics), and for artificial-intelligence researchers automating the extraction of reasoning from natural-language documents.

## References

1. Schum, D.A.: *Evidential Foundations of Probabilistic Reasoning*. John Wiley & Sons, New York, NY (1994)
2. Kadane, J.B., Schum, D.A.: *A Probabilistic Analysis of the Sacco and Vanzetti Evidence*. John Wiley & Sons, New York, NY (1996)
3. Toulmin, S., Rieke, R., Janik, A.: *An Introduction to Reasoning*. Macmillan, New York, NY (1984)
4. Kyburg, Jr., H.E., Teng, C.M.: *Uncertain Inference*. Cambridge University Press, Cambridge, UK (2001)
5. Brachman, R.J., Levesque, H.J.: *Knowledge Representation and Reasoning*. Elsevier, Amsterdam (2004)
6. Brewka, G., Dix, J., Konolige, K.: *Nonmonotonic Reasoning: An Overview*. CSLI Publications, Stanford, CA (1997)
7. Walton, D., Reed, C., Macagno, F.: *Argumentation Schemes*. Cambridge University Press, Cambridge (2008)

8. Walton, D.N.: *Argument Schemes for Presumptive Reasoning*. Lawrence Erlbaum Associates, Mahwah, NJ (1996)
9. Prakken, H.: Analysing reasoning about evidence with formal models of argumentation. *Law, Probability & Risk* 3(1), 33-50 (2004)
10. Prakken, H., Sartor, G.: The three faces of defeasibility in the law. *Ratio Juris* 17(1), 118-39 (2004)
11. Verheij, B.: *Virtual Arguments: On the Design of Argument Assistants for Lawyers and Other Arguers*. TMC Asser Press, The Hague (2005)
12. Prakken, H., Reed, C., Walton, D.: Argumentation schemes and generalisations in reasoning about evidence. In: *International Conference of Artificial Intelligence and Law Proceedings '03*, pp. 32-41. ACM, New York (2003)
13. Gordon, T.F., Walton, D.: Legal Reasoning with Argumentation Schemes. In: *12<sup>th</sup> International Conference on Artificial Intelligence and Law Proceedings '09*, pp. 137-146. ACM, New York (2009)
14. Rissland, E.L., Ashley, K.D., Loui, R.P.: AI and Law: A fruitful synergy. *Artificial Intelligence* 150, 1-15 (2003)
15. Ashley, K.D., Rissland, E.L.: Law, learning and representation. *Artificial Intelligence* 150, 17-58 (2003)
16. Ashley, K.D.: Ontological Requirements for Analogical, Teleological, and Hypothetical Legal Reasoning. In: *12<sup>th</sup> International Conference on Artificial Intelligence and Law Proceedings '09*, pp. 1-10. ACM, New York (2009)
17. Branting, L.K.: *Reasoning with Rules and Precedents: A Computational Model of Legal Analysis*. Kluwer Academic Publishers, Dordrecht (2000)
18. Walker, V.R.: A Default-Logic Paradigm for Legal Fact-Finding. *Jurimetrics* 47, 193-243 (2007)
19. Walker, V.R.: Epistemic and Non-epistemic Aspects of the Factfinding Process in Law. *American Philosophical Association Newsletter on Law and Philosophy* 3(1), 132-136 (2003)
20. Walker, V.R.: Designing Factfinding for Cross-Border Healthcare. *Opinio Juris in Comparatione* 3, paper n. 1, pp. 1-40 (2009)
21. Walker, V.R.: Visualizing the dynamics around the rule-evidence interface in legal reasoning. *Law, Probability and Risk* 6(1 - 4), 5-22 (2007)
22. Walker, V.R.: Emergent Reasoning Structures in Law. In: Trajkovski, G., Collins, S.G. (eds) *Handbook of Research on Agent-Based Societies: Social and Cultural Interactions*, pp. 305-324. Information Science Reference, Hershey, New York (2009)
23. Walker, V.R.: Restoring the Individual Plaintiff to Tort Law by Rejecting "Junk Logic" About Specific Causation. *Alabama Law Review* 56(2), 381-481 (2004)
24. Carmines, E.G., Zeller, R.A.: *Reliability and Validity Assessment*. Sage Publications, Newbury Park (1979)
25. Brewer, S.: Exemplary Reasoning: Semantics, Pragmatics, and the Rational Force of Legal Argument by Analogy. *Harvard Law Review* 109, 923 (1996)