

Logic, Probability Theory and their Application to Legal Reasoning

Soaad Q. Hossain

*Department of Philosophy, Department of Computer and Mathematical Sciences,
University of Toronto, Canada*

Abstract: Both the use of logic and probability theory are heavily applied in many, if not all, areas and industries due to what they can offer. Logic is well known for enabling one to comprehend many things which include intentions, behavior, beliefs, intelligence, knowledge and languages, and to create algorithms designed to solve simple and complex problems. Probability theory is specifically designed to address uncertainty, which uncertainty exists everywhere. Through pairing probability theory with other disciplines is one able to address uncertainty in other fields. Combined, logic and probability theory enable one to address complex problems in law. This paper will first describe logic and probability theory, then address their application to legal reasoning. Accordingly, I argue that logic and probability theory can be used to validate thoughts and arguments made by lawyers and judges, allowing for better understanding of the validity, coherency, truthfulness of their arguments, and the correctness of the statements and decisions made by them.

Keywords: logic, reasoning, legal, probability theory, algorithms, probabilistic, law, belief, evidence, information, sentences

Introduction

From decision-making, thinking and comprehending, to communicating, problem-solving and reasoning, logic and probability theory play a heavy role in everything that we do. For that reason, along with the tools that both can offer, many, if not all, disciplines and industries already or have started using both logic and probability theory to address problems within their respective discipline or industry. With that being said, the field of law can greatly benefit from logic and probability theory. I argue that logic and probability theory can be used to validate thoughts and arguments made by lawyers and judges, allowing for better understanding of the validity, coherency, truthfulness of their arguments, and the correctness of the statements and decisions made by them. I will first describe how logic and probability theory are used, then connect logic and probability theory together and elaborating on how combined together, they can be used to address and approach reasoning. I will then explain how logic and probability theory can be applied to legal reasoning and conclude with elaborating on how they can be used to validate thoughts and arguments used for legal reasoning, and how in doing so can will allow lawyers and judges to bypass lawyers that use psychological weaknesses of humans to their advantage in a way that allows them to construct legal reasoning that stray away from the truth but help them win their cases.

Application of Logic

Within philosophy of language, it has been well documented by numerous philosophers, such as A. J. Ayer, Ludwig Wittgenstein and W.V. Quine, the use of logic in language. For instance, Quine used logic on language to understand language and its relation to reality. In turn, this led him to make many discoveries, which those discoveries have led to advancements within the scientific community. The realization of the role of logic in language motivated certain philosophers to utilize this realization to address challenging questions surrounding intentions, behavior, beliefs, intelligence, and knowledge. One of these questions was “Can machines think?”, which was addressed in Alan Turing’s article *Computing Machinery and Intelligence*. Provided some concepts from psychology and epistemology, Turing used language backed by logic to not only answer the question (his answer being that machines can think), but also defend his answer against possible objections against his answer (Turing, 1950). Fast forwarding to present times, we now see Turing’s answer as the driving force that is positively driving the growth and development of the field of artificial intelligence and machine learning. Accordingly, the development of machine learning algorithms increased from the realization of how efficiently they can solve mathematical and real-world problems.

Logic and Algorithms

An algorithm is essentially a step-by-step procedure for solving a problem. The question that arises is what role logic plays in algorithms. The answer is simple. Within algorithms, their procedures must follow a set of rules, which these rules are defined and communicated through a language. Accordingly, to comprehend those rules, one must understand the language being used to describe the rules, which properly understanding the language requires logic. Through understanding the language and comprehending the rules, individuals, such as computer scientists, can develop algorithms which can then be utilized by themselves or others (e.g. computers) to solve problems. The beauty behind this is that logic enables one to efficiently solve problems given the set of constraints that exists within the problems themselves. However, there are limitations in terms of how much algorithms can do. To understand those limitations would require us to investigate and discuss computability and (computational) complexity of algorithms, which moves away from the purpose of this paper. For that reason, I will not discuss issues and limitations associated with algorithms. The key takeaway here is that logic enables us to approach language, which then combined, logic and language enable us to create algorithms that can be used to solve simple and complex problems; these problems being mathematical and real-world problems, which include behavioral and epistemological problems.

Role of Probability Theory

In solving a problem, probability theory plays a unique role in problem solving. Probability theory enables one to investigate social, psychological, behavioral, physical phenomena that are influenced by randomness and uncertainty. The approach to this involves

Baye's Rule. Baye's Rule provides us with a way to update our belief based on prior information or evidence. More formally, it is defined as

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Where $P(A|B)$ is the probability (known as a conditional probability) that event A occurs given that event B occurred, $P(B|A)$ is the probability (or the likelihood) that event B occurs given that event A occurred, $P(A)$ and $P(B)$ are the probabilities that event A and event B occur, respectively. Accordingly, if we set B as evidence, then $P(A|B)$ would be the probability that event A occurs given evidence B.

What makes probability theory exceptional is that it is a powerful analytical tool which can be search out a set of propositions and detect a contradiction in them if one exists (Jaynes, 2003). This is key to ensuring that each step taken in an algorithm is the correct step taken and needed to optimally solve the problem. Furthermore, this is key to understanding the validity, coherency and truthfulness of statements, such as statements made by lawyers and attorneys. Moving forward, I will explain how probability theory can be used to search out a set of legal propositions and detect a contradiction in them (if one exists), then elaborate on how an algorithm can be created to solve the problem of validating thoughts and arguments used for legal reasoning, which can in turn establish a protocol that can be used to determine the validity, coherency, truthfulness and correctness of legal statements.

Legal Cases and Information

Legal cases are known to be complicated. There is a lot of uncertainty that is involved, making it difficult to make sense of legal propositions and arguments. To obtain validity, coherency and truthfulness of them requires heavily analysis, which unfortunately can be limited based on the time, resources, information, knowledge and mental capability of the individual(s) performing the (legal) analysis. For our cases, we will assume that the adequate the time and resources are provided and those performing the analysis are mentally capable and have the knowledge needed to correctly perform the analysis. However, we will not assume that adequate information is provided as the information provided in each case varies based on the case, evidence provided, etc. For instance, it is possible for false or forged evidence to be presented to the court, which information derived from the evidence would be considered as information, but not lead to adequate information as the quality of the information is not acceptable – such evidence and information derived from that evidence would be considered as false information (purposely) intended to mislead the court. Consequently, what ends up happening is the lawyer and judge are working with less precise information, but are still required to make sense and reason based on that information. To approach this, we can utilize probabilistic logic.

Probabilistic Logic Approach

The probabilistic logic approach is an approach that has been used in artificial intelligence for many years. Nils Nilsson's paper *Probabilistic Logic* described a method that combines

logic with probability theory in a way that probabilistic logical entailment reduces to ordinary logical entailment when the probabilities of all sentences are either 0 or 1. Specifically, in his paper, he presents a semantical generalization of ordinary first-order logic in which the truth values of sentences can range between 0 and 1 (Nilsson, 1986). Of the methods that he mentioned, we will use the one that involves binary semantic tree. Given a set L of sentences, we use a binary semantic tree to determine the sets of consistent truth values. Nilsson describes the method exactly as follows: “*At each node we branch left or right, depending on whether or not we assign one of the sentences in L a value of true or false, respectively. Just below the root we branch on the truth value of one of the sentences in L , next on another sentence in L , and so on. Each path in the tree corresponds to a unique assignment of truth values to the sentences of b ”. We check the consistency of the truth-value assignments as we go, and we close off those paths corresponding to inconsistent valuations. A semantic tree for this example is shown in Figure 1. Closed-off paths are indicated by an x ; consistent sets of valuations are indicated in columns at the tips of their corresponding paths”.*

The example that Nilsson considers in his paper is the sentences $\{P, P \supset Q, Q\}$, and proceeds with constructing the truth table and semantic tree (Nilsson, 1986), shown in Table 1 and Figure 1.

P	<i>true</i>	<i>true</i>	<i>false</i>	<i>false</i>
$P \supset Q$	<i>true</i>	<i>false</i>	<i>true</i>	<i>true</i>
Q	<i>true</i>	<i>false</i>	<i>true</i>	<i>false</i>

Table 1: Truth table for the sentences $\{P, P \supset Q, Q\}$ taken from Nilsson’s paper *Probabilistic Logic*.

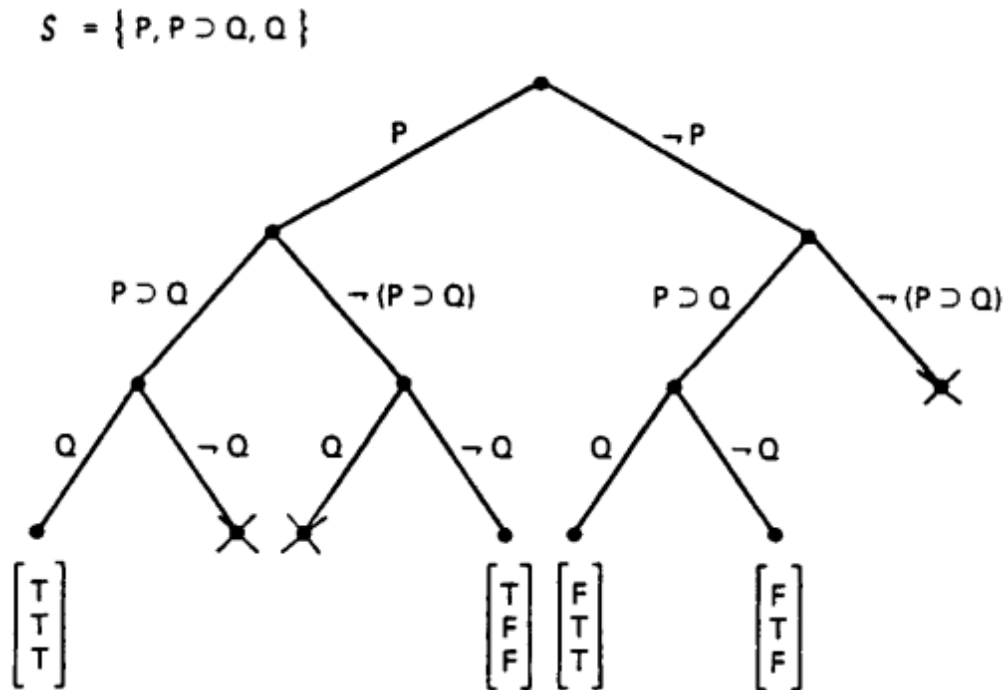


Figure 1: A semantic tree of set S taken from Nilsson’s paper *Probabilistic Logic*. $P \supset Q$ means “P implies Q”, and $\neg P$ and $\neg Q$ are “Not P” and “Not Q” respectively.

Where probability gets involved in this is determining the truth values of sentences P , $P \supset Q$, and Q . That is, determining whether the sentences P is true, $P \supset Q$ is true, and Q is true. Mathematically, these would be denoted as $P(P = 1)$, $P(P \supset Q = 1) = 1$, $P(Q = 1)$, where 1 is equivalent to “true” or “T”.

This leads us to two related types of reasoning problems. The first involves having a base set of sentences (called beliefs) B and its associated probabilities, then using those sentences and the method with probabilistic logic, deduce a new belief S and its associated probability to then solve for P and determine (compute) the probability of S (Nilsson, 1986). The other involves having B and its associated probabilities, then using the method and probabilistic logic, and given that some sentence Q in B is true (or false), we determine the [posterior] probability for some sentence of interest S (Nilsson, 1986). While both types of reasoning problems can involve legal reasoning, we will only address the second problem as that is more closely related to the kind of reasoning used in expert systems, which can be connected to lawyers and judges.

Probabilistic Logic in Legal Reasoning

Within ordinary logic, *modus ponens* allows us to infer Q from P and $P \supset Q$ (Nilsson, 1986) from earlier. When approaching legal reasoning, we must use more than two variables of sentences as there are more variables involved. Let C be denoted as the claims made by the lawyer, A be denoted as the arguments made by the lawyer, S be denoted as the statements made by external individuals (e.g. witnesses, accused), E be denoted as evidence. To evaluate the validity, coherency, truthfulness of arguments made by a lawyer, we: (1) evaluate how truthful E and S ; (2) evaluate the validity of A based on the truthfulness of E and S , (3) evaluate the coherency of C with respect to A , S , and E . This is essentially the way legal reasoning takes place within legal cases. Note that we can use probabilistic logic to denote these as follows:

- (1) $P(E = 1)$ and $P(S = 1)$
- (2) $P(A|E, S)$
- (3) $P(C \supset A \supset (S \wedge E)|C, A, E, S)$

Using probability theory, we can transform each of the above probabilities into activation functions. Let F_i be denoted as the final decisions at moment i , where $i = 1, 2, 3, 4$. Then,

$$\begin{aligned}
 (1) F_1 &= \begin{cases} 1 & \text{if } P(E = 1) \geq \varphi \\ 0 & \text{if } P(E = 1) < \varphi \end{cases} \quad \text{and } F_2 = \begin{cases} 1 & \text{if } P(S = 1) \geq \theta \\ 0 & \text{if } P(S = 1) < \theta \end{cases} \\
 (2) F_3 &= \begin{cases} 1 & \text{if } P(A|E, S) \geq \vartheta \\ 0 & \text{if } P(A|E, S) < \vartheta \end{cases} \\
 (3) F_4 &= \begin{cases} 1 & \text{if } P(C \supset A \supset (S \wedge E)|C, A, E, S) \geq \tau \\ 0 & \text{if } P(C \supset A \supset (S \wedge E)|C, A, E, S) < \tau \end{cases}
 \end{aligned}$$

Where $\varphi, \theta, \vartheta, \tau$ are threshold functions, 1 is denoted as true, and 0 is denoted as false. Translating this into legal perspective, the threshold functions would be the criteria needed to convince the decision-maker (lawyer or judge) that the probability of the thing being presented is convincing enough for them to consider it valid, truthful, or coherent. Notice that different threshold functions are placed within each final decision. The reason for this is

because these thresholds functions heavily vary from one person to another. Furthermore, there are factors that directly or indirectly influence those functions. The decision-maker's mental state (at i), [prior] knowledge, biases, emotional intelligence, mental health, age, experience, and other factors contribute to the threshold functions, which collectively make the values of the thresholds functions a value that is between 0.5 and 1.

In applying the probabilistic logic method from earlier, we can create semantic trees to help us solve for F_1, F_2, F_3 and F_4 . However, since probabilities are involved and we want to address C, A, S and E , we need to utilize Baye's Rule and the knowledge and information that we know and are given. Let K be denoted as our [prior] knowledge and let I be denoted as the information provided about the case (that excludes E and S , but include information such as lawyer profile, mental and behavioral state of witnesses, etc.). Then, when solving for probabilities for (1), we do not solve for $P(E = 1)$ and $P(S = 1)$, but rather we solve for $P(E = 1|K, I)$ and $P(S = 1|K, I)$. Accordingly, as we are given K and I , we use K and I for not only F_1 , but also F_2, F_3 and F_4 . With Baye's Rule being involved, we can use Bayesian Networks from probability theory to approach the Semantic Tree. Bayesian Networks are like semantic trees in the sense of how they are set up; they differ as the variables in Bayesian networks are conditional probabilities.

What is crucial here is that K and I are maximized in the sense that as much knowledge and information is used when approaching each decision. To maximize these, concepts from psycholinguistics, psychology, criminology and other relevant fields should be used. In the case of psycholinguistics, given that it focuses on the relationship between linguistic behavior and psychological processes, it can be used to evaluate S and E . Similarly, other disciplines can be used in the same manner to evaluate S and E . In utilizing them, we can use them as knowledge and information needed to effectively approach legal reasoning with the probabilistic logic method.

Algorithm and Legal Reasoning

Putting everything together, we find that we can create an algorithm that can be used to help us with utilize probability theory and probabilistic logic for validity, coherency, truthfulness within legal reasoning. The algorithm would look something like the following:

- Step 1:** Collect all relevant K and I needed for the case.
- Step 2:** Collect all relevant E and S from lawyers, witnesses, etc.
- Step 3:** Update I accordingly, based on relevant behavior responses from witnesses, etc. and other pertinent details.
- Step 4:** Evaluate the truthfulness of E and S given K and I .
- Step 5:** Evaluate the validity of A given the truthfulness of E and S , and K and I .
- Step 6:** Evaluate the coherency of C with respect to A, S , and E , given A, S, E, K and I .
- Step 7:** Conclude the correctness of the arguments and decisions that was made by lawyers (or judge).

The concluding step of correctness is achieved after obtaining all the details pertaining to the validity, coherency and truthfulness of behind the legal reasoning made by the lawyers or judges.

Closing Remarks

In applying logic and probability theory to legal reasoning, there are several remarks that can be made. The first is that combined, probabilistic logic can be utilized for validity, coherency, truthfulness within legal reasoning, which then probability theory can be used in probabilistic logic to help simplify processes in probabilistic logic. The second remark is that an algorithm can be created to allow one to optimally solve the problem of validity, coherency, truthfulness in claims, statements and arguments made in legal cases. The third and final remark is that in utilizing probabilistic logic in legal reasoning can be very challenging due to the amount of resources, knowledge and information needed to successfully utilize the probabilistic logic method, and the impracticality associated with it. While a software can be created that handles the probabilistic logic component, the software will likely encounter computational complexity issues. Nonetheless, we can always use other methods from probability theory to approach legal reasoning, and logic to create algorithms that can be used to effectively and efficiently address validity, coherency, truthfulness in legal reasoning.

Reference

- Jaynes E. T. (2003). *Probability Theory – The Logic of Science*. Cambridge University Press. 46.
- Nilsson N. J. (1986). Probabilistic Logic. *Artificial Intelligence*. 71 – 87.
- Turing, A. (1950). Computing Machinery and Intelligence. *A Quarterly Review of Psychology and Philosophy*. LIX. 236. 443 – 453.