Reasoning and Knowledge over Impossible Worlds

Alejandro López-Ortiz

Department of Computer Science

University of Waterloo

Waterloo, Ontario, Canada

e-mail: alopez-o@maytag.uwaterloo.ca

Abstract

We study the desired/required abilities of an intelligent agent capable of reasoning about worlds that it knows to be nonexistent. Special attention is given to the Probabilistic Reasoning Model [FHV89,Bac90] and some extensions to this model are proposed and discussed. The notion of independence in this model and the concept of extended probability functions (EPF) are introduced. Independence for assertions that are known to be true or false in all worlds can thus be represented within the EPF framework. The concept of degree of independence is defined as well, and its relationship to the PR model is exemplified.

Keywords: Artificial Intelligence, Probabilistic Reasoning, Knowledge Engineering, Impossible Worlds.

1 Overview

First we present the problem of reasoning in impossible worlds, and explain its relevance and applicability in the real world. Then we determine the model under which the intelligent agent operates. The problem is then recast in this model. In §4 we present a general overview of literature on impossible worlds reasoning or conterfactuals. Then, with this knowledge at hand we define the notion of independence and degrees of independence from which the concept of a measure of dependency follows naturally. We compare this measure with the standard probability of conditionals and Lewis' triviality theorems.

2 Introduction

Reasoning about the real world allows an intelligent agent to emit judgments, make decisions and act [FHV89]. Further, an intelligent agent must be able to reason over imaginary worlds in order to weight the consequences of an event that hasn't happened yet, but that may occur in the future.

However the usefulness of the ability to reason in an imaginary impossible (as opposed to imaginary, but possible) world is not immediate. How would the ability to reason over a world that is known not to exist may benefit an intelligent agent? The following situations exemplify the advantages of such a capability in a variety of situations:

- **Experience**.- In real life, learning through experience represents mostly the concept of making a decision at some point in time and afterwards realizing that there was or was not a better option to be taken at that very moment. Re-evaluating those decisions using the current world information which has been irrevocably modified by reality requires non-standard reasoning capabilities.

- **Gedankenexperiment**.- Scientists in general, and physicists in particular, perform thought experiments or imaginary, impossible experiments whose outcome is of special interest [Ko78]. E.g., suppose a spacecraft travels at the speed of light, then would the passenger of such vehicle stand still in time?

- **Historical Backtracking**.- In order to evaluate the impact of an specific event in time, it is sometimes useful to consider what would have happened if the event in question had not occurred. E.g. assume that Yamamoto had not been killed in the Pacific, would the allied forces then have won the war all the same?

- **Supposition**.- On occasion it is of interest to consider a situation known to be counter to the currently known facts, such as *if Chris were to be the President of the company, the company will go bankrupt.* In this case, we know for a fact that Chris is **not** the president of the company, but still we wish to assume that she is and consider the consequences.[1]

Reasoning over impossible worlds is by no means trivial. A major challenge of such a task is to avoid the undesirable situation where once something non-real has assumed to be true then anything can be assumed to be true. As an example, to assume that a spaceship is traveling at the speed of light should not imply the forfeiture of any other physical law or principle beyond those required for that specific spacecraft to fly at light speed. Moreover, the independence relation is not symmetric, as exemplified by the fact that, while the shape of the continents must not change in the agent's knowledge base (KB) upon any assumption on Hitler's victory, assuming that the

[1] See for example [CH76] and related papers in Counterfactuals.

United Kingdom were not an insular nation would improve Hitler's victory odds in the KB.

For these purposes Chisholm and Stalnaker [Chi46,St68] introduced the so-called Ramsey test which states:

> First, hypothetically make the minimal revision of your stock of beliefs required to assume the antecedent. Then, evaluate the acceptability of the consequent on the basis of this revised body of beliefs.

Stalnaker argues that the condition of minimality is not strictly required. Regardless, once the set of beliefs (i.e. the knowledge base) has been properly modified, the truth value of any statement can be recomputed by the agent, and inferences can be made from it as well.

3 Probabilistic Reasoning (PR)

3.1 Philosophical interpretation of PR

In any probabilistic reasoning model, the agent's degree of belief in any particular instance is represented as a probabilistic value, and as it becomes more cognizant of the surrounding world such degree of belief may change.

Moreover, updates to the knowledge base fall into two categories, depending upon the interpretation of the probabilistic model. Namely, whether the *a priori* probability of an assumption being true is still considered correct or not *after* the deterministic outcome of the experiment.

This can be better understood using an example: suppose that the probability of the Blue Jays winning the pennant has been estimated at 1/2 but afterwards the pennant is won by the Red Soxs. Does that mean that the correct probability of the Blue Jays winning was zero?, or we assume that the probability we assumed is correct and just reality *chooses* randomly between different possibilities?, or perhaps, while we wish to preserve a probabilistic interpretation of the events, the agent may see it fit to update the initial probability estimate, in the light of the subsequent defeat, and then posterior victories in the World Series?

Furthermore, is the probability value of a given statement the expression of the agent's lack of knowledge –and thus of certainty– on the subject or it is representing a truly random phenomena?

At the heart of this matter lies a philosophical question about the interpretation of random phenomena: Are random phenomena truly unpredictable (say a coin toss), or if given enough information, would the agent be able to predict the outcome of the event (e.g. given the initial position, speed, direction of the coin, etc.)?

Since such a question is beyond the scope of this work, we deal only with those cases where probabilities represent the agent's degree of uncertainty or ignorance about an specific event and not a truly random phenomenon. [2]

3.2 The PR model and impossible worlds

The Impossible World Problem is of special interest when the agent uses the Probabilistic Reasoning Model *PR* [Bac89]. In general terms, and agent operating under the *PR* model reasons in a set of worlds with a probability distribution associated on such set. That is to say, the set of "true" facts vary from world to world. Each world has a probability value assigned to it, which can be interpreted as the probability that all the facts in such world are true in the "real" world.

[2]This is equivalent to trying to guess the outcome of a coin toss which has already been tossed, but that happens to be out of sight.

This distribution induces naturally a probability measure over all boolean formulæ, i.e. the probability that a boolean formula is true equals the sum of the probability of the worlds where the boolean formula is true.

A precise definition of an impossible world varies from one reasoning model to another. In the accessibility model [HM85] an impossible world is any inaccessible world. In the *PR* model inaccessible worlds are represented as worlds of null probability. This particular fact leads us to the following definitions:

Definition 1 *We define the following world structure which we use to interpret the formulas of the language of propositional probabilities.*

$$M = \langle O, S, \vartheta, \mu \rangle$$

Where:

1. *O is a set of individuals representing objects of the domain described by the logic.*
2. *S is a set of states for possible worlds.*
3. *ϑ is a function that associates an interpretation of the language with each world. For every $s \in S$, $\vartheta(s)$ is an interpretation that assigns to every object predicate/function symbol a relation of the right arity over O.*
4. *μ is a discrete probability function on S.*

Definition 2 *The truth value assigned to a formula is determined by the following three parameters: the model M, the current world s, and the variable assignment function ν.*

Definition 3 *A world s satisfies a formula α if $(M, s, \nu) \models \alpha$.*

In the PR model the probability of a given world represents the chances of such a world being a model of the *real* world. Therefore if a world has probability zero, it cannot possibly be the real word. More formally,

Definition 4 *Given an intelligent agent acting over $M = \langle O, S, \vartheta, \mu \rangle$, a world $s \in S$ is said to be impossible with respect to the model M if $\mu(s) = 0$.*

Now if a certain assumption is not true in any of the possible models of the real world, then such assumption cannot be part of reality and therefore is imaginary. Again, more formally,

Definition 5 *Given an intelligent agent acting over $M = \langle O, S, \vartheta, \mu \rangle$, an imaginary assumption α is a formula such that all the worlds that satisfy it are impossible.*

An impossible world is well defined since μ is a discrete probability function. If μ is a continuous distribution function then $\mu(s) = 0, \forall s \in S$ and the definition does not generalize to more complex models of probabilistic reasoning. A reasonable extension would be to define impossible worlds in terms of imaginary assumptions.

Definition 6 (Continuous Case) *An assumption α is imaginary if the set of worlds that satisfy it has measure 0, i.e., $\mu(\{s \in S | (M, s, \nu) \models \alpha\}) = 0$.*

Definition 7 (Continuous Case) *A world is impossible if it satisfies an imaginary assumption.*

To simplify the notation, we define $\mathsf{prob}(\alpha) = \mu(\{s \in S | (M, s, \nu) \models \alpha\})$ and we write $\mathsf{cert}(\alpha)$ when $\mathsf{prob}(\alpha) = 1$. Notice that if α is an imaginary assumption then $\mathsf{prob}(\alpha) = 0$, and $\mathsf{cert}(\neg\alpha)$.

As an example, we have that `Hitler won the war` is an imaginary assumption, since, as it is known to the agent that Hitler lost the war, the assertion $\mathsf{cert}(\neg\mathtt{won}(\mathtt{Hitler}))$ is part of the agent's knowledge base, implying $\mathsf{prob}(\mathtt{won}(\mathtt{Hitler})) = 0$ which is precisely the definition of an imaginary assumption.

With this, we have succeeded in formalizing the concept of impossible worlds and imaginary assumptions. Now, it is necessary to extend the intelligent agent's capabilities for it to work in an imaginary world. In the next section we shall study this problem.

4 Projecting reality into impossible worlds

Reasoning over impossible worlds shares some of the aspects and problems of non-monotonic reasoning (see for example [TG81]), since once an imaginary assumption is made, some of the facts previously known to be true, are no longer so. Therefore the truth value of some assertions needs to be changed, but the choice of which assertions to change is, in general, not unique. This problem has been studied by philosophers [Chi46,Go47,Le73,St68,St70], and some key properties of impossible world reasoning have been identified.

As explained in §2, for the agent's reasoning to be useful, it is imperative to make the updated KB reflect the real world. This key observation is known as the Ramsey test:

> First, hypothetically make the minimal revision of your stock of beliefs required to assume the antecedent. Then, evaluate the acceptability of

the consequent on the basis of this revised body of beliefs

Still, Ramsey test does not suffice to identify all revisions. First, because the concept of "minimal revision" is never formally defined, and, secondly, because at times two revisions can be both minimal but one and only one can be true in order to achieve consistency. This conflict was first observed by Goodman [Go47] and it is illustrated by the following two statements, called *counteridenticals*:

`If I were Julius Caesar, I wouldn't be alive in the twentieth century`,
and
`If Julius Caesar were I, he would be alive in the twentieth century`.

Here, once the antecedent is assumed to be true, the KB needs to be corrected to solve the incompatibility between the facts `Julius Caesar is dead` and `I am alive`. But there is no immediate way of deciding whether I should live or die, and in the real world, the "correct" truth value is dependent on the specific purposes for considering such a false assumption.

Regardless of the method used to discriminate among different assumptions that are incompatible, once the decision is made, it can be expressed as a relation in the KB. So it is clear that a method for updating the KB must be developed.

In the next section, we further study the interrelationship of assumptions by means of the concept of independence; this will provides us with the basis for representing relations between impossible assumptions.

5 Independence

5.1 Definition of independence

As it was noted in §2 and §4, reasoning under an imaginary assumption should modify only those statements that *must* be modified for the reasoning to be consistent. More formally, probabilities of independent events should remain the same under unreal assumptions.

The different degrees of independence between assertions are[3]:

- **Absolute Independence**. Assertions are absolutely independent if knowing the truth value of *any* of them gives no information about the truth value of the others. There is absolute independence between the facts `Tweety is a bird` and `John has cancer`.

- **Asymmetrical Independence**. Assertions are asymmetrically independent if knowing the truth of *one* of them gives no information about the truth value of the others. There exists a one side dependence between `Norbert went out hiking` and `It was a warm day`, since Norbert only hikes on warm days, but the day would have been warm independently of the fact that Norbert did not hike.

- **Asymmetrical Dependence**. Analogously, assertions are asymmetrically dependent if the truth value of *one* of them determines the truth value of all of them.

- **Absolute Dependence**. Assertions are absolutely dependent if the truth value of *any* of them determines the truth value of all others.

[3]For a philosophical discussion about degrees of independence, the reader is referred to [San89].

- **Weak Dependence**. Two assertions are weakly dependent if they are not absolutely or asymmetrically dependent or independent. There can be degrees of weak dependence.

The concept of probabilistic logic independence resembles parallel notions of independence in probability theory. Asymmetrical independence in probability theory is denoted by $prob(\beta|\alpha) = \beta$, asymmetrical dependence by $\alpha \rightarrow \beta$, and absolute dependence by $\alpha \equiv \beta$. These definitions do not carry over properly to the domain of impossible worlds. Thus we propose an alternative scheme which matches the probabilistic dependence in worlds of non-null probability.

To express the intuitive notions of in/inter/dependence in the framework of a reasoning model is by no means trivial. Let us explore some approaches.

5.2 Expressing independence

In Axiomatic Theory, a concept α is absolutely independent if the theories having α and $\neg\alpha$ are both logically sound. There are several examples of this notion of independence in mathematics, e.g. the Continuum Hypothesis. Nevertheless this criterion of independence cannot be easily applied to the PR model since its knowledge base is not axiomatic and not axiomatizable in general[4].

Also as it has been pointed out by several authors (e.g. [Fu89]), material implication of standard logic ($p \Rightarrow q \equiv \neg p \wedge q$) does not capture our intuitions about

[4] An axiomatization of a knowledge base would imply a hierarchy in the knowledge base, where some assertions are axioms and some are "theorems" of those axioms. This hierarchy, apart from being artificial, assumes that it is possible to reduce the real world to a set of *independent* axioms and logical consequences of them. Fact is, most things that are known to be true are so simply because they happened to occur and no amount of reasoning would imply with certainty their occurrence (e.g. the Blue Jays are champions because they won, and not as a result of a logic implication).

what implication is. But even if it did, an intuitive implication represents absolute dependence, which is not the negation of absolute independence but its contraposition. Therefore even extended implication concepts as *strong implication* [FHV90] or *relevance logic implication* [RR72] do not model the concept of absolute independence.

For the case of the PR model we have somehow better approximations to a satisfactory definition of absolute independence. Two assertions α and β are *absolutely independent* if $\mathsf{prob}(\alpha|\beta) = \mathsf{prob}(\alpha)$ and$\mathsf{prob}(\beta|\alpha) = \mathsf{prob}(\beta)$ or equivalently α and β are independent (probabilistic independence) if

$$\mathsf{prob}(\alpha \wedge \beta) = \mathsf{prob}(\alpha) \cdot \mathsf{prob}(\beta)$$

Similarly, α is independent of β if $\mathsf{prob}(\alpha|\beta) = \mathsf{prob}(\alpha)$. Note that $\mathsf{prob}(\alpha|\beta)$ is undefined when $\mathsf{prob}(\beta) = 0$. In other words, for imaginary assumptions the PR model cannot describe the concepts of absolute and asymmetrical independence.

Nevertheless, in most cases this notion of independence corresponds to our intuition of independence, save for the special cases when $\mathsf{prob}(\alpha) = 0$ or $\mathsf{prob}(\alpha) = \mathsf{prob}(\beta) = 1$. In this cases, the definition above results in degeneracies[5].

As an example we have that if

$$\mathsf{cert}(\texttt{Blue Jays Won}), \mathsf{cert}(\texttt{Cold war ended}), \mathsf{cert}(\texttt{Gorbachev was president})$$

are assertions of the reasoning model M, then by the probability definition the fact `Blue Jays lost` is absolutely independent of the fact `Blue Jays won` ! since

$$\mathsf{prob}(\texttt{Blue Jays lost}|\texttt{Blue Jays won}) = \mathsf{prob}(\texttt{Blue Jays lost}) = 0.$$

[5]This phenomena resembles the main problem of material implication: $p \Rightarrow q$ does not follows our intuition on implication when either p is false in all worlds (a contradiction) or p and q are true in all worlds (tautologies).

Similarly the end of the cold war is –under this definition– absolutely independent of Gorbachev having been president since

$$\mathsf{prob}(\texttt{Cold war ended}|\texttt{Gorbachev was president}) = \mathsf{cert}(\texttt{Cold war ended}),$$

which contradicts our everyday notion of independence.

In the next section we present some other problems arising from the expression of logical dependence as probabilities.

5.3 Conditional probabilities in philosophy

The notion of conditional probabilities has been thoroughly studied within philosophy of logic circles. Stalnaker, Lewis and Adams have noted key consequences of denoting probabilities of conditionals as conditional probabilities. Scholar research on probabilistic reasoning with conditionals must heed to results obtained in philosophy. It is because of this that we present a summary of philosophical results relevant to the impossible worlds reasoning problem.

In this subsection we present some of the most relevant work in the area and which, in particular, trace paths that can and cannot be traversed without raising local inconsistencies.

Among the problems thus identified are the so called *Triviality Results* [Le76], viz., it is not possible to represent a *closed, logic conditional* dependence between an assumption (imaginary or not) and an assertion with conditional probabilities. We include a proof of this fact to give some insight into the peculiarities of dependence.

Definition 8 *Let $\alpha \rightsquigarrow \beta$ denote the (conditional) dependence relation* "α **is true, then would** $\alpha \Longrightarrow \beta$ **be true?**"*. I.e. $\alpha \rightsquigarrow \beta$ denotes the subset of cases when the implication $\alpha \Longrightarrow \beta$ is true and α is true.*

Definition 9 (St70) *Stalnaker's Hypothesis.-* **Probability of conditionals are conditional probabilities.** *Formally* $\mathsf{prob}(\alpha \rightsquigarrow \beta) = \mathsf{prob}(\beta|\alpha)$.

Definition 10 *An operator* $\oplus$ *such that* $P(A \oplus B) = P(B|A)$*, is closed under conditionalization if* $P(A \oplus B|C) = P(B|AC)$ *for all* A, B, C.

Theorem 1 (Le76) *If* $\rightsquigarrow$ *is closed under conditionalization for a family of probabilities then Stalnaker's Hypothesis is false.*

Proof. See [Le76], where it is shown that if we identify logic and probabilistic dependence, the model is either inconsistent or all probabilities are 0 or 1; thus making the probabilistic model *trivial.*

Therefore, any model of probabilistic expression of dependence cannot be closed under logic conditionalization unless the model is trivial.

5.4 Extensions to the probabilistic model

From the discussion above it follows that a dependence relation that follows our intuitive notion of independence cannot be readily defined.

With this in mind, we first extend the definition of probability measure to probabilities conditioned on imaginary assumptions. This will allow the agent to reason in impossible worlds even in those cases in which dependence remains undefined.

Some authors leave conditional probabilities undefined when the condition event has probability zero [Fe50], while others prefer to assign a zero-one probability function to impossible conditions [It78]. In general, assigning *any* probability distribution to impossible conditional probabilities preserves the validity of most probability theorems. Therefore we have the following definition:

$$
\begin{tikzcd}
\mathcal{A} \times \mathcal{A} \arrow[r, "P_A(\cdot)"] \arrow[d, "P^B(\cdot)"'] \arrow[dr, "\mathcal{P}(B|A)"] & \mathcal{A} \times \mathcal{F} \arrow[d, "P_A(B)"] \\
\mathcal{F} \times \mathcal{A} \arrow[r, "P^B(A)"'] & {[0,1]}
\end{tikzcd}
$$

Figure 1: The Extended Conditional Probability $\mathcal{P}$

Definition 11 *Let $(\Omega, \mathcal{A}, P)$ be a probability space, (i.e. Ω is the set of events, $\mathcal{A}$ is the σ-algebra of measurable sets, and P is the Probability measure) and let A and B be elements of $\mathcal{A}$, where $A \neq \emptyset$. If $P(A) > 0$, the conditional probability of B under A is defined to be equal to $P(A \cap B)/P(A)$ and is denoted by $P(B|A)$ or by $P_A(B)$. If $P(A) = 0$, we define the* extended conditional probability *by $\mathcal{P}(B|A) = P_A(B) = P^B(A)$, where $\mathcal{P} : \mathcal{A} \times \mathcal{A} \longrightarrow \mathcal{A} \times \mathcal{F} \longrightarrow [0,1]$, and $\mathcal{P} : \mathcal{A} \times \mathcal{A} \longrightarrow \mathcal{F} \times \mathcal{A} \longrightarrow [0,1]$ with $\mathcal{F}$ being a set of probability measures (see figure 1).*

Under this definition P_A is a probability function which measures the degree of certainty of a given proposition under the assumption that A is true. Similarly P^B is the probability distribution of our changing belief in B as the current set of worlds changes.

As with standard conditional probabilities, independence between assertions can readily be defined using extended conditional probabilities.

Definition 12 *Let A, B be any two events. Then A is independent of B if $\mathcal{P}(A|B) = P(A)$.*

This extended definition of conditional probability has the same advantages and disadvantages of the original conditional probability with the exception that it applies to imaginary worlds as well. This fact allows to discriminate pathological cases such as $\mathsf{prob}(\texttt{Blue Jays lost}|\texttt{Blue Jays won}) = \mathsf{prob}(\texttt{Blue Jays lost}) = 0$.

In this example, the defeat of the Blue Jays remains independent from their victory but their victory is now totally dependent on their defeat, which before was undefined. That is, $\mathsf{Prob}(\texttt{Blue Jays won}|\texttt{Blue Jays lost}) := 0 \neq \mathsf{prob}(\texttt{Blue Jays won}) = 1$, where Prob is the extended probability operator.

Under this conditions, an intelligent agent now satisfies the conditions of Ramsey's test and can properly update and reason over its KB.

6 Dependence defined

It is possible to define probabilistic dependence as the multiplier that modifies the degree of belief on an assertion. We define the dependence relationship as relative to the probability of the implicant.

Definition 13 *The unnormalized dependence between the truth value of two assertions α and β with respect to the extended probability operator* prob *is given by* $\mathsf{dep}'(\alpha, \beta) = \mathsf{prob}(\alpha|\beta)/\mathsf{prob}(\alpha)$*, for α such that* $\mathsf{prob}(\alpha) > 0$*. The assertions are independent if* $\mathsf{dep}'(\alpha|\beta) = 1$*.*

The dep' function is clearly not a probability measure since it takes values on the interval $(0, \infty)$. For this we define the normalized dependence function.

Definition 14 *The normalized dependence or dependence between two assertions α and β is given by* $\mathsf{dep}(\alpha, \beta) = [\mathsf{dep}'(\alpha, \beta) - 1]/\mathsf{dep}'(\alpha, \beta)$*. Two assertions are independent if* $\mathsf{dep}(\alpha, \beta) = 0$

Notice then that a negative dependence value implies that α is less likely to be true if β is true, while a positive dependence value signifies that α is more likely to be true if β were to be true.

Notice that dep is not a monotone function, and so it follows that it is not subadditive measure (neither is dep' for that matter).

The dep function is rigid (in the sense of [Ba90]) since prob is rigid, this is quite natural since the dependence between two assertions represents the measure of the veracity of an assertion in the set of all the worlds where another is known to be true.

With this definition of dependence we can adjust the modified reality within impossible worlds. Thus, given an impossible assumption β, the agent modifies its knowledge depending upon the value of $\mathsf{dep}(\alpha, \beta)$, for each α in the logic model.

7 Reasoning on impossible worlds

At this point, the model of the impossible world is structurally the same as the model of the real world, so the reasoning mechanisms used by the agent over the KB of the possible worlds can be applied in the KB of the impossible world model. This characteristic is of high relevance, since the agent should reason under the same logic model in either an impossible or a real world.

Once deductions are obtained, the agent still needs to transfer its experience back to the real world KB. Experience can be used to actualize the probabilistic values of assertions. And since the world is non-monotonic, impossible worlds can become possible, e.g. the agent can be certain about `Freedom of Expression`, but a new era of McCarthism will imply a modification to the assertion $\mathsf{prob}(\texttt{Freedom of Expression}) = 1$.

Indeed, it can be argued that non-monotonic and impossible world reasoning are very highly correlated, and that unless certainties are immutable under the non-monotonic model, then impossible world constructions and dependencies are needed.

Among all the reasoning models which have been proposed in the literature, some of them are so general that they cannot be used on today's computers in real time, and others are computationally feasible but too weak to appropriately model reality. Current research is intent on closing the gap between these two approaches. It has been shown that an accessibility relation and a belief operator are required when trying to use first order logic to model the world. The PR model has all this expressiveness, and, if the probability function is extended, reasoning on impossible worlds becomes possible.

As material for further exploration, we could consider the representation of logical impossibility using the extended probability function. So, for sentences which are logically impossible, such as $\alpha \equiv \beta \wedge \neg\beta$ we could have $\mathsf{prob}(\alpha|\alpha) = 0$, contrasted to unreal assumptions that are logically possible for which we have $\mathsf{prob}(\alpha|\alpha) = 1$.

8 Acknowledgements

I would like to thank Fahiem Bacchus, for bringing this problem to my attention and for unvaluable comments on an earlier version of this work.

Bibliography

[**Ba90**] F. Bacchus. *Representing and reasoning with probabilistic knowledge.* The MIT press, Cambridge, 1990.

[**Chi46**] R.M. Chisholm. The Contrary-to-Fact Conditional. *Mind*, Vol. 55, 1946, pp.289-307.

[**CH76**] T.A. Climo, and P.G.A. Howells. Possible worlds in Historical Explanation. *History and Theory*, Vol. 15, 1976, pp.1-20.

[**FHV89**] R. Fagin, J. Y. Halpern, and M. Y. Vardi. *What is an inference rule?*. 1989.

[**Fe50**] W. Feller. *An introduction to probability theory and its applications.* Wiley International Edition, 1950.

[**Fr76**] B. Van Fraasen. Probabilities of Conditionals. *Foundations of probability theory, statistical inference and statistical theories of science*, Vol. 1, 1976, pp.261-308.

[**Fu89**] J. Fulda. Material Implication Revisited. *American Mathematical Monthly*, March, Vol. 96, 1989, pp.247-250.

[**Go47**] N. Goodman. The problem of conterfactual conditionals. *The Journal of Philosophy*, Vol. 44, 1947, pp.113-128.

[**HM85**] J. Y. Halpern and Y. Moses. A guide to modal logics of knowledge and belief. *In Proc. International Joint Conference on Artificial Intelligence (IJCAI)*, 1985, pp.480-490.

[**Ha89**] A. Hájek. Probabilities of Conditionals Revisited. *Journal of Philosophical Logic*, Vol. 18, 1989, pp.423-428.

[**Hi75**] J. Hintikka. Impossible possible worlds vindicated. *Journal of Philosophical Logic*, Vol.44, 1975.

[**It78**] K. Itô. *Introduction to probability theory.* Cambridge University Press, 1984 (Translation from Japanese).

[**Ko78**] G. Koleta. Asking impossible questions about the economy and getting impossible answers. *Science*, Vol. 234, 1985, pp.545-546.

[**Ky90**] H.K. Kyburg, Jr., R.P. Loui, and G.N. Carlson Eds. *Knowledge Representation and Defeasible Reasoning.* Kluwer Academic Publishers, Dordrecht, 1990.

[**Le73**] D. Lewis. Conterfactuals and Comparative Possibility. *Journal of Philosophical Logic*, Vol. 2, 1973, pp.418-446.

[**Le76**] D. Lewis. Probabilities of Conditionals and Conditional Probabilities. *Philosophical Review*, Vol. 85, 1976, pp.297-315.

[**RR72**] R. Routley and V. Routley. Semantics of first degree entailment. *Noûs*, Vol. 6, 1972, pp.335-359.

[**San89**] D.H. Sanford. *If P, then Q*. Routledge, 1989.

[**St68**] R. Stalnaker. A theory of conditionals. *Studies in Logical Theory*, Mon. 2, 1968, pp.98-112.

[**St70**] R. Stalnaker. Probability of Conditionals. *Philosophy of Science*, Vol. 37, 1970, pp.64-80.

[**TG81**] R.H. Thomason, and A. Gupta. A Theory of Conditionals in the Context of Branching Time. In *IFS, Conditionals, Belief, Decision, Chance, and Time.* W.L. Harper, R. Stalnaker, and G. Pearce, Eds. D.Reidel Publishing Company, 1981.