# LOGIC OF PROBABILITY AND CONJECTURE

HARRY CRANE

Abstract. I introduce a formalization of probability in intensional Martin-Löf type theory (MLTT) and homotopy type theory (HoTT) which takes the concept of 'evidence' as primitive in judgments about probability. In parallel to the intuitionistic conception of truth, in which 'proof' is primitive and an assertion $A$ is judged to be true just in case there is a proof witnessing it, here 'evidence' is primitive and $A$ is judged to be probable just in case there is evidence supporting it. To formalize this approach, we regard propositions as types in MLTT and define for any proposition $A$ a corresponding *probability type* $\mathbf{Prob}(A)$ whose inhabitants represent pieces of evidence in favor of $A$. Among several practical motivations for this approach, I focus here on its potential for extending meta-mathematics to include conjecture, in addition to rigorous proof, by regarding a 'conjecture in $A$' as a judgment that '$A$ is probable' on the basis of evidence.

## 1. Introduction

I aim to develop a formal logic of probability and conjecture which takes judgments about evidence as primitive and whose calculus is based on rules for reasoning about such judgments. I focus specifically here on the formal approach and its basic technical consequences. Much more can be said about a wide range of philosophical, historical, and technical motivations underlying this work, including its potential implications for understanding how probability and evidence often arise in legal proceedings, scientific discovery, and everyday decision making. For these additional considerations, I refer the reader to the longer version of this article [Cra18].

The study of conjecture and plausible reasoning in mathematics provides one concrete motivation for the theory introduced here. Although conjecture is central to mathematical practice, and is often a necessary precursor to rigorous proof, there lacks a 'meta-mathematical' framework for analyzing how heuristics and intuitions guide formal 'rigorous' mathematics. But although conjecture and plausible reasoning do not obey strict logical rules, they seem to follow a sound rationale based on some commonly applicable techniques; see, e.g., Pólya's two volume work [Pól54] and Mazur's more recent discussion [Maz12]. For example, let $A$ and $B$ be mathematical propositions. Though it may seem logical for a mathematician who (i) conjectures $A$ and (ii) has proven $A \to B$ (i.e., if $A$ then $B$) to also conjecture $B$, there exists no formal theory to justify the move. We could attempt to formalize this rule by combining the conventional probability calculus (i.e., set $\Pr(Q) = 1$ if $Q$ is true and $\Pr(Q) = 0$ if $\neg Q$ is true) with a Lockean thesis for belief (i.e., conjecture $A$ if '$\Pr(A) \geq t$' for some pre-determined threshold $0 \leq t \leq 1$) [Fol92].[1] In this case, since $A \to B \equiv \neg A \vee B$ has been proven, we have $\Pr(\neg A \vee B) = 1$, and since $A$ is conjectured we must have $\Pr(A) \geq t$ by the Lockean thesis. A routine application of the probability calculus gives $\Pr(\neg A) \leq 1 - t$ and

$$1 = \Pr(\neg A \vee B) \leq \Pr(\neg A) + \Pr(B) \quad \Rightarrow \quad \Pr(B) \geq t,$$

leading to a conjecture in $B$. But whereas in classical logic any proposition $Q$ is either true or false by the law of excluded middle, a conjecture about $Q$ under this approach relies on the extra-mathematical data encoded by the probability operator $\Pr(\cdot)$ and the Lockean threshold $t$.

Here I seek a logic in which truths and conjectures can be treated as mathematical objects of equal standing, thus allowing 'plausible reasoning' in mathematics to be studied internally to the same formal system in which rigorous mathematics already takes place. From this perspective, the logical rules governing 'rigorous mathematics', which is concerned only with truth and proof and not with conjecture and evidence, are a fragment of a more general framework that also incorporates conjecture. I achieve this by introducing a concept of probability on top of the existing syntax of Martin-Löf intensional type theory (MLTT) [ML84, ML87, ML96], and by interpreting a 'conjecture in $A$' as a formal judgment that '$A$ is probable' in the type theory.[2]

### 1.1. Meaning of a conjecture.

Wittgenstein's conception of 'meaning as use' [Wit73] figures into Martin-Löf's meaning explanation of types [ML87, ML96] and the Curry–Howard 'propositions as types' correspondence [Cur34, CF, How69], by which a mathematical proposition $A$ is represented as a type $A : \mathbf{Type}$ whose terms $a : A$ are proofs of $A$:

> "the meaning of a proposition [...] is determined by that which counts as a verification of it." (Martin-Löf [ML96, p. 27])

Following Martin-Löf, we formalize probability in terms of evidence, by associating each proposition $A : \mathbf{Type}$ to a probability type $\mathbf{Prob}(A) : \mathbf{Type}$ whose terms $a' : \mathbf{Prob}(A)$ correspond to pieces of evidence in favor of $A$. In this case, the judgment $a' : \mathbf{Prob}(A)$ indicates that '$a'$ is evidence for $A$' (or that $a'$ witnesses the probability of $A$). In the example discussed above, the conjecture in $A$ corresponds to a judgment $a' : \mathbf{Prob}(A)$, the truth of $A \to B$ corresponds to a proof $f : A \to B$, and the derived conjecture in $B$ results from the formal judgment $\mathtt{imp}_f(a) : \mathbf{Prob}(B)$,[3] which can be constructed by applying the elimination rule (4) for the probability type in the

[1] There are other frameworks for belief revision, some based on the probability calculus and some not. I present this one here only for illustration. See [SF18, AGM85] and references therein for more detailed accounts.

[2] In the specific application to meta-mathematics for conjecture, I interpret the judgment that '$A$ is probable' to correspond to a conjecture in $A$. My natural language use of 'probable' differs from the use of 'plausible' in [Pól54, Maz12] on the grounds that mere 'plausibility' is not sufficient to conjecture $A$. In order to conjecture $A$, one must believe that $A$ is 'likely' or 'probable' on the basis of the observed evidence, not that it is merely 'plausible'.

[3] In MLTT the 'proof' $f : A \to B$ is simply a function which converts any proof of $A$ (i.e., $a : A$) into a proof of $B$ (i.e., $f(a) : B$).

upcoming formalism. Thus, in parallel to Martin-Löf's meaning explanation of truth in terms of proof, we obtain a meaning explanation of probability in terms of evidence:

> the meaning of a conjecture is determined by that which counts as evidence in favor of it.[4]

Before beginning, I note that the intuitionistic approach to probability presented here seems to be autonomous from other conceptions of probability found throughout the literature, including Weatherson's work on 'intuitionistic probability' [Wea03]. In particular, I am not aware of any previous type-theoretic accounts of probability or any other formalism which treats conjectures as first class mathematical objects in their own right. I also note that the concept of 'probability' invoked here is intended in a purely epistemic sense, and thus is best compared and contrasted with other accounts of logical and subjective probability in [Car50, dF, Ram26]. Though not presented in traditional 'theorem-proof' style, the numbered statements in the coming sections can be proven rigorously. To aid the exposition, I defer these proofs to the appendix (Section 5), and opt to explain their meaning as intuitively and pre-formally as possible in terms of probability, evidence, and conjecture. One feature of the constructive logic of MLTT is its amenability to formalization in a computer-based proof assistant, such as Coq or UniMath. Some parts of the formal system introduced here have been formalized in UniMath, and it is left as a topic of future work to fully formalize this theory. Though I attempt to provide as much explanation about type theory as possible, for brevity I assume basic knowledge of the rules and syntax of MLTT. I refer the reader to [Uni13, ML84] for a thorough introduction of MLTT and homotopy type theory. For more on the philosophy of intuitionism, see Brouwer [Brob, Broa, Bro81], Heyting [Hey71], Dummett [Dum00], and Martin-Löf [ML87, ML96].

## 2. Type theoretic probability

At a conceptual level, the state of affairs decomposes statements about truth and probability into five main components:

- **Judgment**: A judgment is an assertion, based on context and witnessed by either proof or evidence, that a concept is true or probable.
- **Context**: Judgments about truth and probability depend on context.
- **Concept**: Judgments are made about concepts. Following Martin-Löf's meaning explanation, the meaning of a concept is determined by how the concept manifests itself.
- **Proof**: The truth of any concept manifests itself in proof.
- **Evidence**: A conjecture about a concept manifests itself in evidence.

[4] Contrast this with the orthodox Bayesian approach to evidence, which takes the probability function $\Pr(\cdot)$ as primitive and defines $E$ as evidence for $A$ just in case $\Pr(A \mid E) > \Pr(A)$, where $\Pr(A \mid E)$ is the conditional probability of $A$ given $E$. From this perspective, the probabilities determined by $\Pr(\cdot)$ exist prior to and independently of judgments about evidence. In the perspective taken here, evidence is primitive: a judgment about the 'probability of $A$' cannot be made without evidence that supports the judgment. I explore this contrast further in [Cra18].

Here I write 'concept' instead of 'proposition' to suggest the more general settings (e.g., law, science) in which this thinking applies. The interpretation of types as mathematical concepts has also been suggested in [LP14]. To emphasize this interpretation below, I write $A$ : **Concept** in place of $A$ : **Type**, without any change to the rules of MLTT.

These main components are expressed formally in MLTT with the following notation.

| pre-formal | formal (MLTT) |
|---|---|
| Context | $\Gamma$ |
| $A$ is a concept | $A : \mathbf{Concept}$ |
| '$A$ is probable' is a concept | $\mathbf{Prob}(A) : \mathbf{Concept}$ |
| judgment that $A$ is true | $a : A$ |
| conjecture that $A$ is true | $a : \mathbf{Prob}(A)$ |

In particular, if $A : \mathbf{Concept}$ is a mathematical proposition, then $a : A$ is interpreted as '$a$ is a proof of $A$' and $a : \mathbf{Prob}(A)$ as '$a$ is evidence for $A$'. I call any judgment '$a : A$' a *truth statement* and '$a : \mathbf{Prob}(A)$' a *conjecture*.

The above setup is formalized in the following syntax, with the left of the turnstile providing *context* for the *judgments* on the right, written 'Context $\vdash$ Judgment'.

*Formal*: $\Gamma \vdash A : \mathbf{Concept}$

*Semi-formal*: Context $\vdash$ $A$ is a concept

*Interpretation*: Context invokes the concept $A$.

*Example*: The context of arithmetic invokes the concept that addition is commutative: for all $n, m : \mathbb{N}$, $n + m = m + n$.

*Formal*: $\Gamma, A : \mathbf{Concept} \vdash \mathbf{Prob}(A) : \mathbf{Concept}$

*Semi-formal*: Context, $A$ is a concept $\vdash$ '$A$ is probable' is a concept

*Interpretation*: The context along with the concept $A$ invokes the concept of the probability of $A$.

*Example*: The context of arithmetic and the concept that every even integer greater than 4 is the sum of two odd primes invokes the concept that such a claim is probable.

*Formal*: $\Gamma, A : \mathbf{Concept} \vdash a : A$

*Semi-formal*: Context, $A$ is a concept $\vdash$ $a$ is a proof of $A$

*Interpretation*: A judgment from context and a concept $A$ that the truth of $A$ is witnessed by proof $a$.

*Example*: A proof from the rules of arithmetic that addition is commutative.

| | |
|---|---|
| *Formal*: | $\Gamma, A : \mathbf{Concept} \quad \vdash a : \mathbf{Prob}(A)$ |
| *Semi-formal*: | Context, $A$ is a concept $\quad \vdash a$ is evidence for $A$ |
| *Interpretation*: | A judgment from context and a concept $A$ that a conjecture in $A$ is supported by evidence $a$. |
| *Example*: | Empirical evidence that each of $6, \ldots, 198, 200$ can be expressed as the sum of two odd primes supports Goldbach's conjecture. |

The interpretation of $\mathbf{Prob}(A)$ as a body of evidence which supports conjectures about $A$ follows by introducing appropriate rules for the probability type.

2.1. **The Probability Type.** Reasoning about concepts in MLTT proceeds by applying rules of formation, introduction, elimination, and computation associated to each type. For lack of space, I refer the reader to the appendix of [KL11] and also [Uni13, Appendix A.2] and [ML84, LP14] for a list of the standard rules of MLTT. See Table 1 for a comparison between the syntax of MLTT and classical logic. Note well the distinction between the constructive, proof relevant syntax of MLTT and the non-constructive, truth functional syntax of classical logic. For example, the type $\sum_{a:A} B(a)$ in MLTT and $\exists a B(a)$ in classical logic can both be read as 'there exists $a$ such that $B(a)$', but their interpretations differ: in MLTT this statement requires an explicit witness $\langle a, b\rangle$ for $a : A$ and $b : B(a)$, and any subsequent outcomes derived from $\langle a, b\rangle : \sum_{a:A} B(a)$ are formally obtained by manipulating $\langle a, b\rangle$ according to the rules of MLTT. In classical logic, on the other hand, the 'truth' of $\exists a B(a)$ is sufficient for deriving further conclusions, without regard for the proof that establishes this claim.

The syntax of MLTT aligns the intuitionistic conceptions of truth and probability through the analogy

$$\text{proof} : \text{truth} \quad :: \quad \text{evidence} : \text{probability}. \tag{1}$$

Whereas the lefthand side of (1) is fulfilled by judgments of the form $a : A$ (in Martin-Löf's meaning explanation), the righthand side is filled out by the following rules for the *probability type*.

Formation rule:

$$\frac{\Gamma \vdash A : \mathbf{Concept}}{\Gamma \vdash \mathbf{Prob}(A) : \mathbf{Concept}} \quad (\mathbf{Prob}\text{-form}) \tag{2}$$

| | |
|---|---|
| *Semi-formal*: | Associated to any concept $A$, whose meaning is determined by proofs that $A$ is true, is a concept $\mathbf{Prob}(A)$, whose meaning is determined by evidence that $A$ is true. |
| *Example*: | From the concept of 'it is raining', derive the concept that 'it is probably raining'. |

Introduction rule:

$$\frac{\Gamma \vdash A : \mathbf{Concept}}{\Gamma, a : A \vdash \mathtt{evid}_A(a) : \mathbf{Prob}(A)} \quad (\mathbf{Prob}\text{-intro}) \tag{3}$$

| Type theory | Interpretation (Curry–Howard) | Classical logic | Interpretation |
|---|---|---|---|
| $A : \mathbf{Type}$ | $A$ is a proposition | $A$ | proposition |
| $a : A$ | $a$ is a proof of $A$ | $\vdash A$ | $A$ is true |
| $A \times B$ | $A$ and $B$ | $A \wedge B$ | $A$ true and $B$ true |
| $A + B$ | $A$ or $B$ (disjoint union) | $A \vee B$ | $A$ true or $B$ true |
| $A \to B$ | $A$ implies $B$ | $\neg A \vee B$ | if $A$ then $B$ |
| $\sum_{a:A} B(a)$ | there exists $a : A$ s.t. $B(a)$ | $\exists a\ B(a)$ | there exists $a$ s.t. $B(a)$ |
| $\prod_{a:A} B(a)$ | $B(a)$ for every $a : A$ | $\forall a\ B(a)$ | $B(a)$ true for all $a$ |
| $\mathbf{0}$ | empty type (contradiction) | $\perp$ | false/contradiction |
| $A \to \mathbf{0}$ | $A$ is false | $\neg A$ | not $A$ |
| $a =_A a'$ | proofs that $a$ and $a'$ are identical | | |

Table 1. Comparison between syntax of MLTT and classical logic. Note that the constructive nature of MLTT means that a proof of '$B(a)$ for all $a : A$' requires an explicit term $f : \prod_{a:A} B(a)$ that associates a witness $f(a) : B(a)$ to every $a : A$. Similarly, a proof that 'there exists $a : A$ such that $B(a)$' requires a witness $a : A$ along with $b(a) : B(a)$ so that $\langle a, b(a)\rangle : \sum_{a:A} B(a)$. The (propositional) identity type $a =_A a'$, which consists of all proofs that $a$ and $a'$ are identical terms of $A$, has no analog in classical logic.

*Semi-formal*: Proof is the definitive and strongest kind of evidence: any proof of $A$ (i.e., $a : A$) determines evidence for $A$ (i.e., $\mathtt{evid}_A(a) : \mathbf{Prob}(A)$).

*Example*: From definitive proof that 'it is raining' (e.g., seeing rain through the window), deduce the weaker statement that 'it is probably raining'.

Elimination rule 1 (Rule of implication):

$$\frac{\begin{array}{c}\Gamma \vdash A : \mathbf{Concept} \\ \Gamma, x : \mathbf{Prob}(A) \vdash C(x) : \mathbf{Concept} \\ \Gamma, a : A \vdash d(a) : C(\mathtt{evid}_A(a))\end{array}}{\Gamma, x : \mathbf{Prob}(A) \vdash \mathtt{imp}_d(x) : \mathbf{Prob}(C(x))} \quad (\mathbf{Prob}\text{-elim-1}) \tag{4}$$

*Semi-formal*: Reasoning about evidence is guided by reasoning about proof. Thus, if $A$ implies $C$, then evidence for $A$ implies evidence for $C$.

*Example*: Since 'it is raining' implies that 'the roads are wet', evidence that 'it is raining' supports the conjecture that 'the roads are probably wet'.

Elimination rule 2 (Rule of inference):

$$\frac{\begin{array}{c}\Gamma \vdash A : \mathbf{Concept} \\ \Gamma, x : \mathbf{Prob}(A) \vdash C(x) : \mathbf{Concept} \\ \Gamma, a : A \vdash d(\mathtt{evid}_A(a)) : C(\mathtt{evid}_A(a))\end{array}}{\Gamma, x : \mathbf{Prob}(A) \vdash \mathtt{inf}_d(x) : C(x)} \quad (\mathbf{Prob}\text{-elim-2}) \tag{5}$$

*Semi-formal*: If $A$ implies $C$ but only through the evidence induced by proofs of $A$, then evidence of $A$ implies $C$.

*Example*: 'It is snowing' implies that 'it is cold outside', but only through the evidence required to determine that there is evidence for 'it is probably snowing'. Thus, from the judgment 'it is probably snowing' we can deduce that 'it is cold outside'.

Computation rule 1:

(6)

$$\frac{\begin{array}{c}\Gamma \vdash A : \mathbf{Concept} \\ \Gamma, x : \mathbf{Prob}(A) \vdash C(x) : \mathbf{Concept} \\ \Gamma, a : A \vdash d(a) : C(\mathtt{evid}_A(a))\end{array}}{\Gamma, a : A \vdash \mathtt{imp}_d(\mathtt{evid}_A(a)) \equiv \mathtt{evid}_{C(\mathtt{evid}(a))}(d(a)) : \mathbf{Prob}(C(\mathtt{evid}_A(a)))} \quad (\mathbf{Prob}\text{-comp-1})$$

*Semi-formal*: The computation rule combines (3) and (4) to make $\mathtt{imp}_d$ defined in (4) compatible with its constructor $d$.

Computation rule 2:

$$\text{(7)} \qquad \frac{\begin{array}{c}\Gamma \vdash A : \mathbf{Concept} \\ \Gamma, x : \mathbf{Prob}(A) \vdash C(x) : \mathbf{Concept} \\ \Gamma, a : A \vdash d(\mathtt{evid}_A(a)) : C(\mathtt{evid}_A(a))\end{array}}{\Gamma, a : A \vdash \mathtt{inf}_d(\mathtt{evid}_A(a)) \equiv d(\mathtt{evid}_A(a)) : C(\mathtt{evid}_A(a))} \quad (\mathbf{Prob}\text{-comp-2})$$

*Semi-formal*: The computation rule combines (3) and (5) to make $\mathtt{inf}_d$ defined in (5) compatible with its constructor $d$.

**0**-rule:

$$\text{(8)} \qquad \frac{}{\Gamma \vdash \mathbf{Prob}(\mathbf{0}) = \mathbf{0} : \mathbf{Concept}} \quad (\mathbf{Prob\text{-}0})$$

*Semi-formal*: Since $\mathbf{0}$ : **Concept** is the concept of emptiness, i.e., has no terms, $\mathbf{Prob}(\mathbf{0})$ is the body of evidence supporting the claim '$\mathbf{0}$ is inhabited'. But since $\mathbf{0}$ is uninhabited by definition, any evidence for $\mathbf{0}$ would immediately lead to contradication; thus, $\mathbf{Prob}(\mathbf{0})$ is also uninhabited.

Additional elimination rules could be introduced to reflect other ways to reason with evidence, e.g., a relation for judging relative strength of evidence (see Section 2.6). To streamline the discussion here, I restrict attention to the fragment described by the rules above.

These rules complete the righthand side of the analogy in (1), warranting the interpretation

$$\begin{array}{ccccccc} a & : & A & :: & \mathtt{evid}_A(a) & : & \mathbf{Prob}(A) \\ \text{proof} & : & \text{truth} & :: & \text{evidence} & : & \text{probability.} \end{array}$$

The next few sections establish some formal consequences of these rules which make precise the pre-formal interpretation of $\mathbf{Prob}(A)$ as a body of evidence and the judgments of the form $a : \mathbf{Prob}(A)$ as conjectures about the truth of $A$. To help the exposition, I defer formal proof of these results to Section 5.

2.2. **Carriers of probability.** By the introduction rule (3), evidence can be derived from any proof, and thus to every $a : A$ there corresponds a piece of evidence $\mathtt{evid}_A(a) : \mathbf{Prob}(A)$. But because judgments about probability are conjectural, the interpretation of the probability type in terms of evidence is viable only if the logic permits the judgment $a : \mathbf{Prob}(A)$ to be made without definitive proof of $A$. In other words, the formal calculus should be consistent with but should not require

$$\prod_{a:\mathbf{Prob}(A)} \sum_{x:A} (\mathtt{evid}_A(x) =_{\mathbf{Prob}(A)} a). \tag{9}$$

> *To every piece of evidence for $A$ (i.e., $a : \mathbf{Prob}(A)$) there exists a proof of $A$ (i.e., $x : A$) that is compatible with that evidence (i.e., $p : \mathtt{evid}_A(x) = a$).*

If, for example, (9) were required to hold, then any conjecture $a : \mathbf{Prob}(A)$ would have to correspond to a proof, thus defeating the purpose of the established formalism as a logic for handling evidence. But while it is possible that a conjecture $a : \mathbf{Prob}(A)$ may be valid without any $x : A$ for which $\mathtt{evid}_A(x) = a$, the formalism is consistent with this stringent correspondence between $\mathbf{Prob}(A)$ and $A$:

$$\prod_{a:\mathbf{Prob}(A)} \left( \left( \sum_{x:A} (\mathtt{evid}_A(x) =_{\mathbf{Prob}(A)} a) \to \mathbf{0} \right) \to \mathbf{0} \right). \tag{10}$$

> *It cannot be ruled out that every conjecture (i.e., $a : \mathbf{Prob}(A)$) corresponds to a proof (i.e., $x : A$) through $\mathtt{evid}_A : A \to \mathbf{Prob}(A)$.*

According to (10), the logic is compatible with judgments of someone who refuses to conjecture without having definite proof.

The observation in (10) goes hand in hand with the intended interpretation of the inhabitants of $\mathbf{Prob}(A)$ as carriers of evidence. For if

$$\sum_{x:A} (\mathtt{evid}_A(x) =_{\mathbf{Prob}(A)} a) \to \mathbf{0}$$

were consistent for some $A$ and $a : \mathbf{Prob}(A)$, then the interpretation of $a$ as evidence for $A$ would be called into question. Since

$$\sum_{x:A} (\mathtt{evid}_A(x) =_{\mathbf{Prob}(A)} a) \to \mathbf{0}$$

rules out that $a$ corresponds to some proof of $A$, the interpretation of $a$ as evidence supporting the conjecture that $A$ is provable becomes obscured. So even though the logic is consistent with probability judgments which do not necessarily correspond to a direct proof, in order for a probability judgment ($a : \mathbf{Prob}(A)$) to qualify as a credible statement about 'evidence in favor of $A$', *bona fide* evidence $a : \mathbf{Prob}(A)$

must at least *suggest* that $a$ corresponds to a proof of $A$:

$$\prod_{a:\mathbf{Prob}(A)} \mathbf{Prob}\left(\sum_{x:A} (\mathtt{evid}_A(x) =_{\mathbf{Prob}(A)} a)\right). \tag{11}$$

> *Any piece of evidence for $A$ (i.e., $a : \mathbf{Prob}(A)$) gives rise to evidence that there exists a proof of $A$ (i.e., $x : A$) that is compatible with that evidence (i.e., $p : \mathtt{evid}_A(x) = a$).*

2.3. **Structure of probability.** The structure of probability induced by the formal rules (2)-(8) can be summarized neatly for free-standing concepts (i.e., non-dependent types). Suppose $A, B : \mathbf{Concept}$ is such that $B$ is provable whenever $A$ is, i.e., $A \to B$ is inhabited. Informally, since the truth of $A$ implies the truth of $B$ (i.e., $A \to B$) and evidence for $A$ hints at the truth of $A$, then evidence for $A$ should also hint at the truth of $B$. Formally, this follows by applying the elimination rule (4) to any $f : A \to B$. In particular, define $C \equiv \lambda a.B : A \to \mathbf{Concept}$ in (4)[5] so that for any $f : A \to B$ and any context $\Gamma$ one can derive[6]

$$\Gamma,\ a : A \vdash f(a) : C(a) \equiv B.$$

Then by (4) and (6), we have $\mathtt{imp}_f : \mathbf{Prob}(A) \to \mathbf{Prob}(B)$ such that $\mathtt{imp}_f(\mathtt{evid}_A(a)) \equiv \mathtt{evid}_B(f(a))$. When combined with (5) and (7), we obtain the following commutative diagram for $A, B, C : \mathbf{Concept}$.

$$\begin{tikzcd}
A \arrow[r, bend left, "f"] \arrow[r, bend right, "d \circ \mathtt{evid}_A"'] \arrow[rr, bend left=40, "g \circ f"] \arrow[d, "\mathtt{evid}_A"] & B \arrow[r, bend left, "g"] \arrow[r, bend right, "e \circ \mathtt{evid}_B"'] \arrow[d, "\mathtt{evid}_B"] & C \arrow[d, "\mathtt{evid}_C"] \\
\mathbf{Prob}(A) \arrow[r, "\mathtt{imp}_f"'] \arrow[ur, "\mathtt{inf}_d"'] \arrow[rr, bend right=40, "\mathtt{imp}_g \circ \mathtt{imp}_f = \mathtt{imp}_{g \circ f}"'] & \mathbf{Prob}(B) \arrow[r, "\mathtt{imp}_g"'] \arrow[ur, "\mathtt{inf}_e"'] & \mathbf{Prob}(C)
\end{tikzcd} \tag{12}$$

[5]Here I have used $\lambda$-abstraction to define the function $C : A \to \mathbf{Concept}$ which assigns each $a$ to $B$. In general the notation $\lambda x.y : X \to Y$ is a function that assigns each $x : X$ to some $y : Y$.

[6]Given $a : A$ and $f : A \to B$, we can apply $f$ to $a$ to obtain $f(a) : B$.

$$\mathtt{imp}_f \circ \mathtt{evid}_A \equiv \mathtt{evid}_B \circ f : A \to \mathbf{Prob}(B)$$
$$\mathtt{inf}_d \circ \mathtt{evid}_A \equiv d \circ \mathtt{evid}_A : A \to B$$

$$\mathtt{imp}_{d \circ \mathtt{evid}_A} = \mathtt{imp}_{\mathtt{inf}_d \circ \mathtt{evid}_A} = \mathtt{inf}_{\mathtt{evid}_B \circ \mathtt{inf}_d} \tag{13}$$

$$\mathtt{imp}_{e \circ \mathtt{evid}_B} = \mathtt{imp}_{\mathtt{inf}_e \circ \mathtt{evid}_B} = \mathtt{inf}_{\mathtt{evid}_C \circ \mathtt{inf}_e}. \tag{14}$$

Note that for general expressions $x$ and $x'$ of type $A$, I write $x \equiv x' : A$ to denote *judgmental equality* and $x =_A x'$ to denote *propositional identity* in MLTT. Propositional identity $x =_A x'$ is shorthand for the *type* $\mathtt{Id}(x, x', A)$ in MLTT, whose terms are proofs that $x$ and $x'$ are identical. When the type is clear from context, I often omit it, writing $x = x'$ for propositional identity.

2.4. **Combining evidence.** The rules of MLTT together with the rules for the probability type induce a logic for combining evidence of different assertions and deriving new conjectures from old. For example, as carriers of evidence for '$A$ and $B$', the terms of $\mathbf{Prob}(A \times B)$ ought to give rise to evidence for $A$ and evidence for $B$ individually. That is, when considering the conjecture $x : \mathbf{Prob}(A \times B)$ for '$A$ and $B$', one can disregard the relevance of $x$ to $B$ (respectively, $A$) and derive a conjecture for $A$ (resp. $B$) on its own:

$$\mathtt{splitprob}_{A,B} : \mathbf{Prob}(A \times B) \to \mathbf{Prob}(A) \times \mathbf{Prob}(B). \tag{15}$$

*Evidence for '$A$ and $B$' can be split into separate pieces of evidence for $A$ and $B$ individually.*

Similarly, when in possession of evidence for $A$ or evidence for $B$, one can derive evidence for '$A$ or $B$':

$$\mathtt{combprob}_{A,B} : \mathbf{Prob}(A) + \mathbf{Prob}(B) \to \mathbf{Prob}(A + B). \tag{16}$$

*Evidence for '$A$ or $B$' can be derived from evidence for $A$ or evidence for $B$.*

Together (15) and (16) give the hierarchy

$$\mathbf{Prob}(A \times B) \to \mathbf{Prob}(A) \times \mathbf{Prob}(B) \to \mathbf{Prob}(A) + \mathbf{Prob}(B) \to \mathbf{Prob}(A + B). \tag{17}$$

The implications in (17) do not, in general, go in reverse. One may not, for example, feel compelled to combine two pieces of evidence, one for $A$ and one for $B$, into a single conjecture for $A$ and $B$ jointly, as it might not be clear whether the specific pieces of evidence, say, $a : \mathbf{Prob}(A)$ and $b : \mathbf{Prob}(B)$, are compatible as evidence for '$A$ and $B$'. Similarly, having evidence for '$A$ or $B$' need not be sufficient for deriving evidence for either of the two individually. Indeed, for a given proposition $A$, one might postulate the law of excluded middle $\mathtt{LEM}_A : A + \neg A$ without specifying which of $A$ or $\neg A$ holds. Asserting $\mathtt{LEM}_A$ allows the derivation $\mathtt{evid}_{A+\neg A}(\mathtt{LEM}_A) : \mathbf{Prob}(A + \neg A)$ by (3), but without further evidence as to which of $A$ or $\neg A$ holds the logical calculus does not justify a conjecture in $A$ or $\neg A$ individually. These observations illustrate the 'proof relevant' character of MLTT: e.g., to conjecture '$A$ and $B$', it is not enough to simply have evidence for $A$ and evidence for $B$; the two pieces of evidence must also be compatible with one another.

The relations shown in (12), (15), and (16) for non-dependent types $A$ and $B$ extend to relations about universal and existential statements for dependent types $B : A \to \mathbf{Concept}$.

$$\mathbf{Prob}(\prod_{a:A} B(a)) \to \prod_{a:A} \mathbf{Prob}(B(a)). \tag{18}$$

*From evidence that $B$ holds for every proof of $A$, derive evidence of $B(a)$ from any particular proof $a : A$.*

$$\sum_{a:A} \mathbf{Prob}(B(a)) \to \mathbf{Prob}(\sum_{a:A} B(a)). \tag{19}$$

*From any proof of $A$ (i.e., $a : A$) for which there is evidence for $B$ (i.e., $b : \mathbf{Prob}(B(a))$), derive evidence that $B$ holds for some proof of $A$ (i.e., $\mathbf{Prob}(\sum_{a:A} B(a))$).*

2.5. **Logic for handling evidence.** The opening discussion of Section 1 raises the question of how a conjecture for $B$ can be justified from (i) a conjecture for $A$ and (ii) proof of $A \to B$. In our formalism, this corresponds to exhibiting a term of type

$$\mathbf{Prob}(A) \times (A \to B) \to \mathbf{Prob}(B),$$

which we have already shown through the commutative diagram in (12). Other similar inference rules for the probability type can be derived from the rules (2)-(8) by noting the distinction between $A \to B$ in MLTT and classical logic. In classical logic, $A \to B$ is defined as the material conditional $\neg A \vee B$, read as 'if $A$ then $B$'. By the law of excluded middle this is equivalent to $\neg\neg(\neg A \vee B) \equiv \neg(A \wedge \neg B)$, yielding logical equivalence among the statements

$$\neg A \vee B \equiv A \to B \equiv \neg(A \wedge \neg B).$$

But in the constructive logic of MLTT, the above three statements have different meanings and are related by the hierarchy

$$(\neg A + B) \to (A \to B) \to \neg(A \times \neg B), \tag{20}$$

which in turn elicits a corresponding hierarchy among the corresponding probability statements; see (27).

2.5.1. *Evidence under $\neg A + B$.* Working from left to right in (20), assume first that there is evidence for $A$ and proof of '$B$ or not $A$' (i.e., $\neg A + B$). By the elimination rule for the coproduct type $\neg A + B$, we reason by case analysis. If $B$ is already known, then $B$ can trivially be derived, regardless of the evidence for $A$. And by the implication rule (4) and $\mathbf{0}$-rule (8), $\neg A \equiv A \to \mathbf{0}$ implies $\mathbf{Prob}(A) \to \mathbf{0}$, meaning that the evidence for $A$ can be used to derive a contradiction, from which anything follows. Together this gives an inhabitant of

$$\mathbf{Prob}(A) \times (\neg A + B) \to B. \tag{21}$$

*From a conjecture in $A$ and proof of '$B$ or not $A$', derive a proof of $B$.*

Notice as a special case that taking $B \equiv A$ in (21) gives $\mathbf{Prob}(A) \times (A + \neg A) \to A$, so that, in particular, evidence for $A$ and $\mathtt{LEM}_A : A + \neg A$ is enough to construct a proof of $A$. Thus, from evidence of $A$ one can derive truth of $A$ deductively by postulating $\mathtt{LEM}_A : A + \neg A$. Though this seems counterintuitive at first, it is clarified by considering the meaning of the judgment $\mathtt{LEM}_A : A + \neg A$ in the logic of MLTT. (In MLTT, $\mathtt{LEM}_A : A + \neg A$ is a 'proof' that '$A$ or $\neg A$' holds. By the introduction rule for coproducts, such a proof corresponds either to a proof of $A$ or a proof of $\neg A$. If the former, then $A$ is true. If the latter, then the evidence for $A$ contradicts the proof of 'not $A$', and from a contradiction anything follows.)

On the other hand, from proof of $A$ and a conjecture '$B$ or not $A$, one derives evidence for $B$:

$$A \times \mathbf{Prob}(\neg A + B) \to \mathbf{Prob}(B). \tag{22}$$

2.5.2. *Evidence under $A \to B$.* The structure of evidence summarized in (12) implies that evidence for $B$ can be derived from evidence for $A$ and proof that '$A$ implies $B$', i.e., from $a : \mathbf{Prob}(A)$ and $f : A \to B$ derive $\mathtt{imp}_f(a) : \mathbf{Prob}(B)$. In fact, this can be done so that the evidence for $A$ is compatible with the derived evidence for $B$, justifying a conjecture in '$A$ and $B$' jointly:

$$\mathbf{Prob}(A) \times (A \to B) \to \mathbf{Prob}(A \times B) \tag{23}$$

> *Given evidence for $A$ and proof that '$A$ implies $B$', conjecture '$A$ and $B$'.*

Also by (4), any verification of $A$ and evidence that '$A$ implies $B$' combine into evidence for '$A$ and $B$':

$$A \times \mathbf{Prob}(A \to B) \to \mathbf{Prob}(A \times B). \tag{24}$$

2.5.3. *Evidence under $\neg(A \times \neg B)$.* When equipped with proof that 'A and not $B$ is not the case' (i.e., $\neg(A \times \neg B)$), evidence for $A$ must suggest that $\neg B$ is not the case (i.e., $\neg\neg B$) since the evidence for $A$ hints that $A$ is true while the proof of $\neg(A \times \neg B)$ implies that $A$ and $\neg B$ cannot both be true:

$$\mathbf{Prob}(A) \times \neg(A \times \neg B) \to \mathbf{Prob}(A \times \neg\neg B). \tag{25}$$

> *Evidence for $A$ and proof of $\neg(A \times \neg B)$ combine to give evidence for $A \times \neg\neg B$.*

(Recall that, in general, $\neg\neg B$ and $B$ are not identical in MLTT, because MLTT does not require the law of excluded middle.)

Similarly, from proof of $A$ and evidence for $\neg(A \times \neg B)$, one can derive evidence for $\neg\neg B$ since knowing $A$ and having evidence that $A$ and $\neg B$ cannot both hold gives evidence against $\neg B$:

$$A \times \mathbf{Prob}(\neg(A \times \neg B)) \to \mathbf{Prob}(A \times \neg\neg B). \tag{26}$$

Comparing the statements in (21), (23), and (25) and ignoring the appearance of $A$ in their conclusion[7] gives the following commutative diagram, which agrees with the hierarchy in (20).

(27)

$$\begin{array}{ccc} \neg A + B & \qquad & \mathbf{Prob}(A) \times (\neg A + B) \longrightarrow B \\ \downarrow & & \downarrow \qquad\qquad\qquad \downarrow \\ A \to B & & \mathbf{Prob}(A) \times (A \to B) \longrightarrow \mathbf{Prob}(B) \\ \downarrow & & \downarrow \qquad\qquad\qquad \downarrow \\ \neg(A \times \neg B) & & \mathbf{Prob}(A) \times \neg(A \times \neg B) \longrightarrow \mathbf{Prob}(\neg\neg B) \end{array}$$

2.6. **Grades of evidence.** We have so far discussed the probability type as a way of representing evidence 'at level one', i.e., evidence of a proposition. Unlike the conventional numerical approach to probability as a measure of evidence, the formalism presented here provides no immediate way to compare different pieces of evidence in terms of which is 'stronger'. In the same way that different proofs of $A$ cannot be compared on the basis of which better establishes the truth of $A$—both establish the truth of $A$, but in (possibly) different ways—there is no aspect of the formal system which allows for one to judge, for example, that $a : \mathbf{Prob}(A)$ is 'stronger evidence' for $A$ than $a' : \mathbf{Prob}(A)$, or that $a$ makes $A$ 'more probable' than $a'$ does.

There are two possible ways to incorporate such a notion of evidential strength into this formalism. One is to extend upon the rules (2)-(8) by adding a relation $\leq_A$: $\mathbf{Prob}(A) \times \mathbf{Prob}(A) \to \mathbf{Bool}$ for each $A : \mathbf{Concept}$ along with rules for $\leq_A$ that agree with the interpretation of $\leq_A (a, a')$ as '$a'$ is stronger evidence for $A$ than $a$'. A second possibility requires no additional rules, but instead follows by iterating the probability type constructor $\mathbf{Prob} : \mathbf{Concept} \to \mathbf{Concept}$ to obtain an inductive hierarchy of different 'grades of evidence', beginning with truth ($A$), then evidence of truth ($\mathbf{Prob}(A)$), evidence of evidence ($\mathbf{Prob}(\mathbf{Prob}(A))$), evidence of evidence of evidence ($\mathbf{Prob}(\mathbf{Prob}(\mathbf{Prob}(A)))$), and so on. For each $A : \mathbf{Concept}$, the formalism captures these different grades of evidence by the inductive definition

$$\begin{aligned} \mathbf{Prob}_n(A) &: \mathbf{Concept} \\ \mathbf{Prob}_0(A) &:\equiv A \\ \mathbf{Prob}_{n+1}(A) &:\equiv \mathbf{Prob}(\mathbf{Prob}_n(A)), \end{aligned}$$

so that $\mathbf{Prob}_n(A)$ consists of the *nth level evidence of* $A$. The theorems expressed throughout Section 2.5 can be extended to these 'higher probability types' in a

[7]For example, from the conclusion $\mathbf{Prob}(A \times \neg\neg B)$ in (26), we can apply (15) to get $\texttt{splitprob}_{A,\neg\neg B} : \mathbf{Prob}(A \times \neg\neg B) \to \mathbf{Prob}(A) \times \mathbf{Prob}(\neg\neg B)$, which we compose with the projection map $\texttt{pr}_{\neg\neg B} : \mathbf{Prob}(A) \times \mathbf{Prob}(\neg\neg B) \to \mathbf{Prob}(\neg\neg B)$, $\langle x, y\rangle \mapsto y$, to obtain $\mathbf{Prob}(\neg\neg B)$.

straightforward way. For example, for $m, n \geq 0$, we can extend (23) to

$$\mathbf{Prob}_m(A) \times \mathbf{Prob}_n(A \to B) \to \mathbf{Prob}_{m+n}(B).$$

Both of these possible extensions are interesting for future examination, but are not discussed here due to space limitations; see [Cra18] for further discussion.

## 3. Homotopy Probability Theory

So far the logic of probability described above has been purely syntactic, expressed as a logic for reasoning about concepts in MLTT. Though Martin-Löf's meaning explanation provides an interpretation of this syntax in terms of proofs and evidence for propositions, we can gain additional insights by interpreting the syntax of MLTT into homotopy type theory (HoTT). In HoTT, the types in MLTT are interpreted as homotopy types (i.e., topological spaces up to homotopy equivalence), and the calculus is empowered by the resulting univalence axiom, by which two types $A, B$ : **Concept** are regarded as identical (i.e., $A =_{\mathbf{Concept}} B$) if their associated homotopy types are homotopy equivalent.

To emphasize the interpretation in HoTT with univalence, we write $A : \mathcal{U}$ (instead of $A$ : **Concept**) to indicate that $A$ is a homotopy type in a univalent universe $\mathcal{U}$ (i.e., a universe of types for which the univalence axiom holds). With $A \simeq B$ denoting that $A, B : \mathcal{U}$ are *homotopy equivalent*[8] and $A =_{\mathcal{U}} B$ signifying that $A$ and $B$ are *identical* as homotopy types in $\mathcal{U}$, the univalence axiom states, roughly, that for all $A, B : \mathcal{U}$

$$\begin{array}{ccc} (A \simeq B) & \simeq & (A =_{\mathcal{U}} B) \\ \text{equivalence} & \text{is equivalent to} & \text{identity.} \end{array} \tag{28}$$

Formally, this is accomplished by constructing the canonical map stating that identity implies equivalence,

$$\texttt{idtoequiv}_{A,B} : (A =_{\mathcal{U}} B) \to (A \simeq B),$$

and asserting as an axiom,

$$\texttt{ua} : \prod_{A,B:\mathcal{U}} \texttt{isequiv}(\texttt{idtoequiv}_{A,B}),$$

so that $\texttt{idtoequiv}_{A,B}$ is an equivalence between $A$ and $B$ for all $A, B : \mathcal{U}$.[9]

The univalence axiom is a powerful and intriguing proposal in Voevovdsky's Univalent Foundations program [APW13, PW12, Uni13]. As it is impossible to discuss the many interesting aspects of HoTT and univalence in this short presentation, I

---

[8] $A \simeq B$ is formally defined as the type $\sum_{f:A\to B} \texttt{isequiv}(f)$, where

$$\texttt{isequiv}(f) :\equiv \left(\sum_{g:B\to A} f \circ g \sim \texttt{id}_B\right) \times \left(\sum_{h:B\to A} h \circ f \sim \texttt{id}_A\right),$$

for $\texttt{id}_C \equiv \lambda c.c : C \to C$ for any $C : \mathcal{U}$.

[9] Following the convention of [Uni13], I write $\texttt{ua}(p) : A =_{\mathcal{U}} B$ to indicate the image of $p : A \simeq B$ under $\texttt{ua}$.

provide here only a cursory overview of how the familiar probabilistic concepts of independence, conditional probability, and additivity can be conceived in the proposed type theoretic version of probability. I refer the reader to [Awo14, Awo16, APW13, Tse16, Tse17, Shu17, Uni13] for further details about HoTT.

3.1. **Identical concepts have identical probabilities.** It is intuitive that identical concepts ought to have identical probabilities, as can be proven using the induction rule for identity types in MLTT without appealing to univalence: for all $A, B$ : **Concept** and $a, a' : A$,

$$(a =_A a') \to (\mathtt{evid}_A(a) =_{\mathbf{Prob}(A)} \mathtt{evid}_A(a')) \tag{29}$$

$$(A =_{\mathbf{Concept}} B) \to (\mathbf{Prob}(A) =_{\mathbf{Concept}} \mathbf{Prob}(B)). \tag{30}$$

In MLTT, however, there is no mechanism for proving, e.g., $A \times B =_{\mathbf{Concept}} B \times A$. (In MLTT, we can only prove that $A \times B$ and $B \times A$ are isomorphic, i.e., $A \times B \simeq B \times A$, but we need the univalence axiom to derive $A \times B =_{\mathcal{U}} B \times A$ from this isomorphism.) With the univalence axiom, we have, for all $A, B : \mathcal{U}$,

$$(A \simeq B) \to (\mathbf{Prob}(A) \simeq \mathbf{Prob}(B)), \tag{31}$$

from which several obvious statements follow as corollaries, including

$$\mathbf{Prob}(A \times B) \simeq \mathbf{Prob}(B \times A) \quad \text{and}$$
$$\mathbf{Prob}(A + B) \simeq \mathbf{Prob}(B + A).$$

The univalence axiom has additional consequences for more nuanced aspects of the probability type, such as conditional probability, independence, and additivity, which warrant a much more in depth discussion than the brief introduction below.

3.2. **Conditional probability.** In practice, it is common to form a judgment, such as *'A and B' is probable*, by combining new evidence for $B$ with old evidence for $A$. Formally, we define the *conditional probability of $B$ given $a$* : $\mathbf{Prob}(A)$ as the dependent type $\mathbf{Prob}(B \mid -) : \mathbf{Prob}(A) \to \mathbf{Concept}$ given for each $a : \mathbf{Prob}(A)$ by

$$\mathbf{Prob}(B \mid a) :\equiv \sum_{x:\mathbf{Prob}(A \times B)} (\mathtt{imp}_{\mathtt{pr}_A}(x) =_{\mathbf{Prob}(A)} a). \tag{32}$$

> *Conditional evidence for $B$ given evidence for $A$ (i.e., $a$ : $\mathbf{Prob}(A)$) consists of evidence for 'A and B' (i.e., $x : \mathbf{Prob}(A \times B)$) along with proof that $x$ is compatible with $a$ (i.e., $p : \mathtt{imp}_{pr_A}(x) = a$).*

In (32), $\mathtt{pr}_A : A \times B \to A$, $\langle a, b\rangle \mapsto a$, is the canonical projection map and $\mathtt{imp}_{\mathtt{pr}_A}$ : $\mathbf{Prob}(A \times B) \to \mathbf{Prob}(A)$ is the map constructed by applying the implication rule (4) to $\mathtt{pr}_A$. According to (32), the inhabitants of $\mathbf{Prob}(B \mid a)$ correspond to 'conditional conjectures', i.e., a conjecture in $B$ that is conditional on the conjecture $a : \mathbf{Prob}(A)$. Such a conjecture can be asserted just in case there is evidence $x : \mathbf{Prob}(A \times B)$ for $A$ and $B$ along with proof that $x$ is compatible with $a$. A conditional conjecture in $B$ given $a : \mathbf{Prob}(A)$ is thus stronger than a conjecture in '$A$ and $B$' alone because the conditional probability requires that $x : \mathbf{Prob}(A \times B)$ is compatible with a specific conjecture $a$ : $\mathbf{Prob}(A)$. This may at first seem counterintuitive since it appears that the previous evidence for $A$ has already done "half the work" in establishing the conjecture in '$A$ and $B$'. But, at the same time, $a : \mathbf{Prob}(A)$ constrains what can serve as evidence for the conditional conjecture in $B$, because the conditional conjecture is required to be compatible with $a : \mathbf{Prob}(A)$.

By analogy to the classical law of total probability in the probability calculus,[10] we observe a similar equivalence between $\mathbf{Prob}(A \times B)$ and the total space of all conditional probabilities $\mathbf{Prob}(B \mid a)$ over all $a : \mathbf{Prob}(A)$,

$$\mathbf{Prob}(A \times B) \simeq \sum_{a:\mathbf{Prob}(A)} \mathbf{Prob}(B \mid a). \tag{33}$$

> *Having evidence for 'A and B' is equivalent to having a piece of evidence for A (i.e., $a : \mathbf{Prob}(A)$) along with conditional evidence for B that is compatible with a.*

For any $A, B : \mathcal{U}$, the equivalence (33) is established by the conditionalization map

$$\mathtt{cond}_{B|A} : \mathbf{Prob}(A \times B) \to \sum_{a:\mathbf{Prob}(A)} \mathbf{Prob}(B \mid a)$$

$$\mathtt{cond}_{B|A}(x) :\equiv \langle \mathtt{imp}_{\mathtt{pr}_A}(x), \langle x, \mathtt{refl}_{\mathtt{imp}_{\mathtt{pr}_A}(x)} \rangle \rangle, \tag{34}$$

which decomposes evidence $x : \mathbf{Prob}(A \times B)$ for $A$ and $B$ into evidence for $A$ (i.e., $\mathtt{imp}_{\mathtt{pr}_A}(x) : \mathbf{Prob}(A)$) and conditional evidence for $B$ given $\mathtt{imp}_{\mathtt{pr}_A}(x)$ (i.e., $\langle x, \mathtt{refl}_{\mathtt{imp}_{\mathtt{pr}_A}(x)} \rangle : \mathbf{Prob}(B \mid \mathtt{imp}_{\mathtt{pr}_A}(x))$). I prove (33) in Section 5.

3.3. **Independence.** Two concepts can be regarded as *independent* whenever their associated 'bodies of evidence' are unrelated to one another. In other words, two assertions are independent if a conjecture about one is irrelevant to a conjecture about the other, expressed formally as

$$\mathbf{Prob}(A \times B) =_{\mathbf{Concept}} \mathbf{Prob}(A) \times \mathbf{Prob}(B). \tag{35}$$

> *Evidence for A (resp. B) serves neither to corroborate nor refute evidence for B (resp. A).*

I note in passing the similarity between (35) and the definition of independence in the ordinary probability calculus, $\Pr(A \wedge B) = \Pr(A) \times \Pr(B)$, with $A$ and $B$ regarded as propositions and '$\times$' interpreted as multiplication.

Under univalence, with concepts interpreted as spaces in a universe $\mathcal{U}$, (35) is equivalent to

$$\mathbf{Prob}(A \times B) \simeq \mathbf{Prob}(A) \times \mathbf{Prob}(B), \tag{36}$$

which must be witnessed by a homotopy equivalence between $\mathbf{Prob}(A \times B)$ and $\mathbf{Prob}(A) \times \mathbf{Prob}(B)$. From Section 2.4, there is a canonical mapping

$$\mathtt{splitprob}_{A,B} : \mathbf{Prob}(A \times B) \to \mathbf{Prob}(A) \times \mathbf{Prob}(B)$$

[10] In the standard (numerical) probability calculus, the law of total probability states that $\Pr(A \wedge B) = \sum_{j=1}^{k} \Pr(B \mid A_j)\Pr(A_j)$ for any propositions $A, B$ and a partition of $A$ into mutually exclusive propositions $A \equiv A_1 \vee \cdots \vee A_k$.

as defined in (15), but there is no canonical mapping $\mathbf{Prob}(A) \times \mathbf{Prob}(B) \to \mathbf{Prob}(A \times B)$. By understanding independence to mean that evidence for $A$ is irrelevant to evidence for $B$, and vice versa, we define the independence type of $A$ and $B$ by

$$\texttt{indep}(A, B) :\equiv \texttt{isequiv}(\texttt{splitprob}_{A,B}). \tag{37}$$

3.4. **Conditional independence.** Combining the definitions in Sections 3.2 and 3.3, we define *conditional independence* by replacing the probabilities in (35) with conditional probabilities: for $A, B : \mathbf{Concept}$, we say that $A$ and $B$ are *conditionally independent given* $C : \mathbf{Concept}$ if

$$\prod_{c:\mathbf{Prob}(C)} (\mathbf{Prob}(A \times B \mid c) =_{\mathbf{Concept}} \mathbf{Prob}(A \mid c) \times \mathbf{Prob}(B \mid c)).$$

Interpreting MLTT into HoTT, we formally define the *conditional independence type of $A$ and $B$ given $C$* by

$$\texttt{indep}(A, B \mid C) :\equiv \prod_{c:\mathbf{Prob}(C)} \texttt{isequiv}(\texttt{condsplit}_{A,B|C}(c)),$$

for

$$\texttt{condsplit}_{A,B|C} : \prod_{c:\mathbf{Prob}(C)} \mathbf{Prob}(A \times B \mid c) \to \mathbf{Prob}(A \mid c) \times \mathbf{Prob}(B \mid c)$$

$$\begin{aligned}&\texttt{condsplit}_{A,B|C} :\equiv \\ &\quad \equiv \lambda c.\lambda\langle x, p\rangle.\langle\langle \texttt{imp}_{\texttt{pr}_A}(x), \texttt{comp}_{\texttt{pr}_C, \texttt{pr}_{A\times C}} \bullet p\rangle, \langle \texttt{imp}_{\texttt{pr}_B}(x), \texttt{comp}_{\texttt{pr}_C, \texttt{pr}_{B\times C}} \bullet p\rangle\rangle,\end{aligned}$$

where, for general $f : A \to B$ and $g : B \to C$,

$$\texttt{comp}_{g,f} : \texttt{imp}_g \circ \texttt{imp}_f = \texttt{imp}_{g\circ f}$$

is the proof of the propositional identity for composition reflected in (12), and for general $x, y, z : X$, $r \bullet s : x = z$ is the concatenation of the paths determined by $r : x = y$, and $s : y = z$ in the homotopic interpretation.

3.5. **Additivity.** I conclude with a brief discussion of additivity, which figures prominently in the axioms of conventional probability theory but whose analog is absent from the evidence-based theory presented above. In the ordinary quantitative theory of probability, the additivity axiom says that

$$\Pr(A \vee B) = \Pr(A) + \Pr(B) \tag{38}$$

for any mutually exclusive propositions $A$ and $B$. If these probabilities are interpreted as a measure of the amount of evidence supporting '$A$ and $B$', then (38) says that the amount of evidence supporting '$A$ or $B$' equals the amount supporting $A$ plus the amount supporting $B$. In the type theoretic ('proof relevant') setting, with $\mathbf{Prob}(A)$ and $\mathbf{Prob}(B)$ interpreted as the bodies of evidence supporting $A$ and $B$, respectively, we express the analog to (38) as

$$\mathbf{Prob}(A + B) =_{\mathbf{Concept}} \mathbf{Prob}(A) + \mathbf{Prob}(B), \tag{39}$$

with $+$ interpreted now as the coproduct in type theory. In (39), the lefthand side is the body of evidence for '$A$ or $B$' while the righthand side is the disjoint union of the body of evidence for $A$ and the body of evidence for $B$. The discussion in Section 2.4 showed that evidence for $A$ or evidence for $B$ gives evidence for '$A$ or $B$', i.e.,

$$\texttt{combprob}_{A,B} : \mathbf{Prob}(A) + \mathbf{Prob}(B) \to \mathbf{Prob}(A + B), \tag{40}$$

but in general having evidence for $A$ or $B$ is not enough to determine which of the two there is evidence for.

These observations provide a link between our conception of probability as a body of evidence and the Dempster–Shafer axioms of belief functions $\mathrm{Bel}(\cdot)$ [Dem67, Sha76], which instead of (38) require the weaker condition

$$\mathrm{Bel}(A \vee B) \geq \mathrm{Bel}(A) + \mathrm{Bel}(B). \tag{41}$$

The inequality in (41) reflects the possibility that the amount of evidence favoring $A \vee B$ might strictly exceed the sum of evidence for $A$ and $B$ individually. Note that, by interpreting '$\to$' as '$\leq$' when evidence is regarded as a quantity, the implication in (40) agrees with the inequality in (41). The theory of evidence presented here is thus consistent with the Shaferian mathematical theory of evidence [Sha76]. It is interesting to consider the implications of assuming the additivity condition (39) when interpreted in HoTT. In this case, (39) becomes

$$\mathbf{Prob}(A + B) \simeq \mathbf{Prob}(A) + \mathbf{Prob}(B),$$

and one could postulate (perhaps as an axiom) that the canonical map $\texttt{combprob}_{A,B}$ in (40) is an equivalence,

$$\texttt{isequiv}(\texttt{combprob}_{A,B}).$$

But this is beyond the scope of our discussion here.

## 4. Concluding remarks

I have proposed a type-theoretic formalization of probability in which probability statements are defined as primitive judgments about evidence. As the concepts of probability and evidence have been intermingled for millenia, cf. [GZP+89, Fra15], the formalism presented here is perhaps more historically accurate than the current mathematical orthodoxy for probability. Indeed, it was not until relatively recently in history that probability took its present numerical form [Por96, Hac75]. Also, since judgments of the form '$a$ is evidence for $A$' arise much more commonly and naturally than precise quantitative probability assignments (i.e., degrees of belief), this framework is arguably better for modeling the way in which people routinely reason with evidence in legal proceedings, scientific investigation, mathematical conjecture, and everyday decision making. Finally, I have posed mathematical conjecture as the backdrop in order to anchor the exposition in something concrete without delving too far into the details of the given application. I discuss many more historical, philosophical, and conceptual aspects of this work in [Cra18].

## 5. Appendix: Technical Proofs

*Proof of* (11). We apply the elimination rule (4) to construct a witness

$$\lambda a.\mathtt{imp}_d(a) : \prod_{a:\mathbf{Prob}(A)} \mathbf{Prob}(C(a)),$$

for the type family

$$C : \mathbf{Prob}(A) \to \mathbf{Concept}$$
$$C(a) :\equiv \sum_{x:A}(\mathtt{imp}_{\mathtt{pr}_A}(x) =_{\mathbf{Prob}(A)} a)$$

and $d : \prod_{x:A} C(\mathtt{evid}_A(x))$ defined by

$$d(x) :\equiv \langle x, \mathtt{refl}_{\mathtt{evid}_A(x)}\rangle : \sum_{x:A}(\mathtt{evid}_A(x) =_{\mathbf{Prob}(A)} \mathtt{evid}_A(x)). \tag{42}$$

□

*Proof of* (10). Given $a : \mathbf{Prob}(A)$ and $f : \sum_{x:A}(\mathtt{evid}_A(x) =_{\mathbf{Prob}(A)} a) \to \mathbf{0}$, we can derive $f(y) : \mathbf{0}$ for each $y : \sum_{x:A}(\mathtt{evid}_A(x) =_{\mathbf{Prob}(A)} a)$. By the implication rule (4), we immediately have

$$\mathtt{imp}_f : \mathbf{Prob}\left(\sum_{x:A}(\mathtt{evid}_A(x) =_{\mathbf{Prob}(A)} a)\right) \to \mathbf{Prob}(\mathbf{0}) \equiv \mathbf{0}$$
$$\mathtt{imp}_f(\mathtt{evid}(y)) :\equiv \mathtt{evid}_{\mathbf{0}}(f(y)), \quad y : \sum_{x:A}(\mathtt{evid}_A(x) =_{\mathbf{Prob}(A)} a).$$

Finally, let $d$ be as defined in (42), so that by (11), we have

$$\lambda a.\mathtt{imp}_d(a) : \prod_{a:\mathbf{Prob}(A)} \mathbf{Prob}\left(\sum_{x:A}(\mathtt{evid}_A(x) =_{\mathbf{Prob}(A)} a)\right).$$

We conclude by constructing

$$\lambda a.\lambda f.\mathtt{imp}_f(\mathtt{imp}_d(a)) : \prod_{a:\mathbf{Prob}(A)} (\sum_{x:A}(\mathtt{evid}_A(x) =_{\mathbf{Prob}(A)} a) \to \mathbf{0}) \to \mathbf{0}\,.$$

□

*Proof of* (12), (13), *and* (14). Several of the commutativity relations in (12) follow directly from the rules (2)-(8) of the probability type. For example, by (4) and (6), we immediately have

$$\mathtt{imp}_f \circ \mathtt{evid}_A \equiv \mathtt{evid}_B \circ f : A \to \mathbf{Prob}(B) \quad \text{and}$$
$$\mathtt{inf}_d \circ \mathtt{evid}_A \equiv d \circ \mathtt{evid}_A : A \to B,$$

and analogously for $g : B \to C$ and $e \circ \mathtt{evid}_B : B \to C$ in (12). By the first judgmental equality, we thus have $\mathtt{comp}_{g,f} : \mathtt{imp}_g \circ \mathtt{imp}_f = \mathtt{imp}_{g\circ f}$ by first proving

$$prod_{a:\mathbf{Prob}(A)}(\mathtt{imp}_{g\circ f}(a) =_{\mathbf{Prob}(C)} (\mathtt{imp}_g \circ \mathtt{imp}_f)(a)) \tag{43}$$

and then applying the axiom of function extensionality. To prove (43), we apply the second elimination and computation rules of the probability type ((5) and (7)) as follows. First define

$$C : \mathbf{Prob}(A) \to \mathbf{Concept}$$
$$C(a) :\equiv \mathtt{imp}_{g\circ f}(a) =_{\mathbf{Prob}(C)} (\mathtt{imp}_g \circ \mathtt{imp}_f)(a).$$

For every $x : A$, we have

$$\mathtt{imp}_{g\circ f}(\mathtt{evid}_A(x)) \equiv \mathtt{evid}_C(g(f(x))) \quad \text{and}$$
$$(\mathtt{imp}_g \circ \mathtt{imp}_f)(\mathtt{evid}_A(x)) \equiv \mathtt{imp}_g(\mathtt{imp}_f(\mathtt{evid}_A(x))) \equiv \mathtt{imp}_g(\mathtt{evid}_B(f(x))) \equiv \mathtt{evid}_C(g(f(x))),$$

so that

$$d(\mathtt{evid}_A(x)) \equiv \mathtt{refl}_{\mathtt{evid}_C(g(f(x)))} : C(\mathtt{evid}_A(x))$$

depends on $x$ only through $\mathtt{evid}_A(x)$. By (5), we have

$$\mathtt{inf}_d : \prod_{a:\mathbf{Prob}(A)} C(a),$$

as desired. Commutativity of the other relations in (12) follow by similar applications of the eliminations rules for **Prob**.

To show the first equality in (13), we use both computation rules (6) and (7) with

$$C : \mathbf{Prob}(A) \to \mathbf{Concept}$$
$$C(a) :\equiv \mathtt{imp}_{d\circ \mathtt{evid}_A}(a) = \mathtt{imp}_{\mathtt{inf}_d \circ \mathtt{evid}_A}(a)$$

as follows. For $x : A$, we have

$$\mathtt{imp}_{d\circ \mathtt{evid}_A}(\mathtt{evid}_A(x)) \equiv \mathtt{evid}_B(d(\mathtt{evid}_A(x))) : A \to \mathbf{Prob}(B)$$
$$\mathtt{imp}_{\mathtt{inf}_d\circ \mathtt{evid}_A}(\mathtt{evid}_A(x)) \equiv \mathtt{evid}_B(\mathtt{inf}_d(\mathtt{evid}_A(x))) \equiv \mathtt{evid}_B(d(\mathtt{evid}_A(x))) : A \to \mathbf{Prob}(B),$$

so that

$$r(\mathtt{evid}_A(x)) \equiv \mathtt{refl}_{\mathtt{evid}_B(d(\mathtt{evid}_A(x)))} : C(\mathtt{evid}_A(x))$$

and

$$\mathtt{inf}_r : \prod_{a:\mathbf{Prob}(A)} \mathtt{imp}_{d\circ \mathtt{evid}_A}(a) = \mathtt{imp}_{\mathtt{inf}_d\circ \mathtt{evid}_A}(a)$$

by (7). For the second equality, we argue similarly by noting that for every $x : A$

$$\mathtt{inf}_{\mathtt{evid}_B\circ \mathtt{inf}_d}(\mathtt{evid}_A(x)) \equiv \mathtt{evid}_B(\mathtt{inf}_d(\mathtt{evid}_A(x))) \equiv \mathtt{evid}(d(\mathtt{evid}_A(x))) : A \to \mathbf{Prob}(B)$$

by (7).

□

*Proof of* (15). For $A, B : \mathbf{Concept}$ let

$$\mathtt{pr}_A : A \times B \to A$$
$$\mathtt{pr}_A(\langle a, b\rangle) :\equiv a$$

and

$$\mathtt{pr}_B : A \times B \to B$$
$$\mathtt{pr}_B(\langle a, b\rangle) :\equiv b$$

be the projection maps. By (12) we have

$$\begin{aligned}\mathtt{imp}_{\mathtt{pr}_A} &: \mathbf{Prob}(A \times B) \to \mathbf{Prob}(A) \quad \text{and}\\ \mathtt{imp}_{\mathtt{pr}_B} &: \mathbf{Prob}(A \times B) \to \mathbf{Prob}(B),\end{aligned}$$

from which we construct

$$\lambda x.\langle \mathtt{imp}_{\mathtt{pr}_A}(x), \mathtt{imp}_{\mathtt{pr}_B}(x)\rangle : \mathbf{Prob}(A \times B) \to \mathbf{Prob}(A) \times \mathbf{Prob}(B).$$

□

*Proof of* (16). For $A, B$ : **Concept** let $\mathtt{inl} : A \to A + B$ and $\mathtt{inr} : B \to A + B$ be the left and right injections, respectively. By (12) (cf. (4)) we have

$$\begin{aligned}\mathtt{imp}_{\mathtt{inl}} &: \mathbf{Prob}(A) \to \mathbf{Prob}(A + B) \quad \text{and}\\ \mathtt{imp}_{\mathtt{inr}} &: \mathbf{Prob}(B) \to \mathbf{Prob}(A + B),\end{aligned}$$

from which we construct

$$\begin{gathered}h : \mathbf{Prob}(A) + \mathbf{Prob}(B) \to \mathbf{Prob}(A + B)\\ h(\mathtt{inl}(a)) :\equiv \mathtt{imp}_{\mathtt{inl}}(a), \quad a : \mathbf{Prob}(A),\\ h(\mathtt{inr}(b)) :\equiv \mathtt{imp}_{\mathtt{inr}}(b), \quad b : \mathbf{Prob}(B).\end{gathered}$$

□

*Proof of* (18). Let $B : A \to$ **Concept** be a dependent type. Fix $a : A$ and define $C : \mathbf{Prob}(\prod_{y:A} B(y)) \to \mathbf{Concept}$ as the non-dependent type $C(x) :\equiv B(a)$. We construct $\lambda f.f(a) : \prod_{y:A} B(y) \to B(a)$ so that the elimination rule (4) implies

$$\lambda x.\mathtt{imp}_{\lambda f.f(a)}(x) : \mathbf{Prob}(\prod_{y:A} B(y)) \to \mathbf{Prob}(B(a)).$$

We may thus construct

$$\lambda x.\lambda a.\mathtt{imp}_{\lambda f.f(a)}(x) : \mathbf{Prob}(\prod_{y:A} B(y)) \to \prod_{a:A} \mathbf{Prob}(B(a)).$$

□

*Proof of* (19). Let $B : A \to$ **Concept** be a dependent type. For each $a : A$ define $C_a : \mathbf{Prob}(B(a)) \to \mathbf{Concept}$ as the non-dependent type $C_a(x) :\equiv \sum_{y:A} B(y)$. From any $b : B(a)$ we construct $\langle a, b\rangle : C_a(\mathtt{evid}_{B(a)}(b))$, so that the implication rule (4) implies

$$\lambda x.\mathtt{imp}_{\lambda b:B(a).\langle a,b\rangle}(x) : \mathbf{Prob}(B(a)) \to \mathbf{Prob}(\sum_{y:A} B(y)).$$

We then define

$$\begin{gathered}h : \sum_{a:A} \mathbf{Prob}(B(a)) \to \mathbf{Prob}(\sum_{a:A} B(a))\\ h(\langle a, x\rangle) :\equiv \mathtt{imp}_{\lambda b:B(a).\langle a,b\rangle}(x).\end{gathered}$$

□

*Proof of* (21). Fix $A, B$ : **Concept** and note first that from any $f : \neg A \equiv A \to \mathbf{0}$ the elimination rule (4) implies $\mathtt{imp}_f : \mathbf{Prob}(A) \to \mathbf{0}$. Arguing by case analysis for the coproduct type, we thus construct

$$h : \mathbf{Prob}(A) \times (\neg A + B) \to B$$
$$h(\langle a, \mathtt{inl}(f)\rangle) :\equiv \mathtt{efq}_B(\mathtt{imp}_f(a)), \quad a : \mathbf{Prob}(A),\ f : A \to \mathbf{0},$$
$$h(\langle a, \mathtt{inr}(b)\rangle) :\equiv b, \quad a : \mathbf{Prob}(A),\ b : B,$$

where $\mathtt{efq}_B : \mathbf{0} \to B$ is *ex falso quodlibet* for $B$. □

*Proof of* (22). For $A, B$ : **Concept** and $a : A$, we define $d_a : \neg A + B \to B$ by the elimination rule for $\neg A + B$:

$$d_a(\mathtt{inl}(f)) :\equiv \mathtt{efq}_B(f(a)), \quad f : A \to \mathbf{0},$$
$$d_a(\mathtt{inr}(b)) :\equiv b, \quad b : B.$$

By the elimination rule for the probability type (4), we construct $\mathtt{imp}_{d_a} : \mathbf{Prob}(\neg A + B) \to \mathbf{Prob}(B)$, from which we conclude by defining

$$\lambda a.\lambda x.\mathtt{imp}_{d_a}(x) : A \times \mathbf{Prob}(\neg A + B) \to \mathbf{Prob}(B).$$

□

*Proof of* (23). For $A, B$ : **Concept** and $f : A \to B$, we construct

$$h_f : \mathbf{Prob}(A) \to \mathbf{Prob}(A \times B)$$
$$h_f(a) :\equiv \mathtt{imp}_{\lambda x.\langle x, f(x)\rangle : A \to (A \times B)}(a).$$

We then define

$$\lambda a.\lambda f.h_f(a) : \mathbf{Prob}(A) \times (A \to B) \to \mathbf{Prob}(A \times B).$$

□

*Proof of* (24). For $A, B$ : **Concept** and $a : A$, we define

$$d_a \equiv \mathtt{imp}_{\lambda f.\langle a, f(a)\rangle : (A \to B) \to (A \times B)} : \mathbf{Prob}(A \to B) \to \mathbf{Prob}(A \times B)$$

by the elimination rule (4). We then construct

$$\lambda a.\lambda x.d_a(x) : A \times \mathbf{Prob}(A \to B) \to \mathbf{Prob}(A \times B).$$

□

*Proof of* (25). For $A, B$ : **Concept**, $f : A \times (B \to \mathbf{0}) \to \mathbf{0}$, and $a : A$, we define

$$f_a :\equiv \lambda b.f(a, b) : (B \to \mathbf{0}) \to \mathbf{0}\,.$$

Thus, for every $a : A$ we have $f_a : \neg\neg B \equiv (B \to \mathbf{0}) \to \mathbf{0}$ and $\lambda a.\langle a, f_a\rangle : A \to (A \times \neg\neg B)$. The elimination rule for the probability type (4) gives

$$\mathtt{imp}_{\lambda a.\langle a, f_a\rangle} : \mathbf{Prob}(A) \to \mathbf{Prob}(A \times \neg\neg B).$$

From the judgmental equality

$$A \times (B \to \mathbf{0}) \to \mathbf{0} \equiv \neg(A \times \neg B),$$

we define

$$\lambda x.\lambda f.\mathtt{imp}_{\lambda a.\langle a,f_a\rangle}(x) : \mathbf{Prob}(A) \times \neg(A \times \neg B) \to \mathbf{Prob}(A \times \neg\neg B).$$

□

*Proof of* (26). For $A, B : \mathbf{Concept}$, $a : A$, and $f : A \times (B \to \mathbf{0}) \to \mathbf{0}$, we define

$$f_a :\equiv \lambda b.f(a,b) : (B \to \mathbf{0}) \to \mathbf{0}$$

as in the proof of (25). We then define

$$d_a \equiv \lambda f.\langle a, f_a\rangle : \neg(A \times \neg B) \to (A \times \neg\neg B)$$

so that $\mathtt{imp}_{d_a} : \mathbf{Prob}(\neg(A \times \neg B)) \to \mathbf{Prob}(A \times \neg\neg B)$. The proof is completed by

$$\lambda a.\lambda x.\mathtt{imp}_{d_a}(x) : A \times \mathbf{Prob}(\neg(A \times \neg B)) \to \mathbf{Prob}(A \times \neg\neg B).$$

□

*Proof of* (27). The following commutes:

$$\begin{tikzcd}
\mathbf{Prob}(A) \times (\neg A + B) \arrow[r, "\alpha"] \arrow[d, "\beta"] & B \arrow[d, "\mathtt{evid}_B"] \\
\mathbf{Prob}(A) \times (A \to B) \arrow[r, "\zeta"] \arrow[d, "\gamma"] & \mathbf{Prob}(B) \arrow[d, "\epsilon"] \\
\mathbf{Prob}(A) \times \neg(A \times \neg B) \arrow[r, "\delta"] & \mathbf{Prob}(\neg\neg B)
\end{tikzcd} \tag{44}$$

Define $\alpha, \beta, \gamma, \delta, \epsilon$ as follows. (For $f : A \to C$ and $g : B \to C$, I write $\mathtt{ind}_{f,g} : A + B \to C$ for the function defined by case analysis.)

$$\lambda a.\alpha_L(a) \equiv \lambda a.\lambda g.\mathtt{efq}_B(\mathtt{imp}_g(a)) : \mathbf{Prob}(A) \to \neg A \to B$$
$$\lambda a.\alpha_R(a) \equiv \lambda a.\lambda b.b : \mathbf{Prob}(A) \to B \to B$$
$$\alpha \equiv \lambda a.\lambda z.\mathtt{ind}_{\alpha_L(a),\alpha_R(a)}(z) : \mathbf{Prob}(A) \times (\neg A + B) \to B$$
$$\lambda a.\beta_L(a) \equiv \lambda a.\lambda g.\langle a, \lambda x.\mathtt{efq}_B(g(x))\rangle : \mathbf{Prob}(A) \to \neg A \to \mathbf{Prob}(A) \times (A \to B)$$
$$\lambda a.\beta_R(a) \equiv \lambda a.\lambda b.\langle a, \lambda x.b\rangle : \mathbf{Prob}(A) \to B \to \mathbf{Prob}(A) \times (A \to B)$$
$$\beta \equiv \lambda a.\lambda z.\mathtt{ind}_{\beta_L(a),\beta_R(a)}(z) : \mathbf{Prob}(A) \to (\neg A + B) \to \mathbf{Prob}(A) \times (A \to B)$$
$$\gamma \equiv \lambda a.\lambda f.\langle a, \lambda x.\lambda g.g(f(x))\rangle : \mathbf{Prob}(A) \to (A \to B) \to \mathbf{Prob}(A) \times \neg(A \times \neg B)$$
$$\delta \equiv \lambda a.\lambda g.\mathtt{imp}_g(a) : \mathbf{Prob}(A) \to (A \to \neg\neg B) \to \mathbf{Prob}(\neg\neg B)$$
$$\epsilon \equiv \lambda b.\mathtt{imp}_{\lambda y.\lambda g.g(y):B\to\neg B\to\mathbf{0}}(b) : \mathbf{Prob}(B) \to \mathbf{Prob}(\neg\neg B)$$
$$\zeta \equiv \lambda a.\lambda f.\mathtt{imp}_f(a) : \mathbf{Prob}(A) \times (A \to B) \to \mathbf{Prob}(B).$$

We first show that the upper square commutes by repeated application of the elimination rule for the product, coproduct, and probability types. For the upper half of the square, $\mathtt{evid}_B \circ \alpha : \mathbf{Prob}(A) \times (\neg A + B) \to \mathbf{Prob}(B)$ is defined by case analysis:

$$(a, \mathtt{inl}(g)) \mapsto \mathtt{evid}_B(\mathtt{efq}_B(\mathtt{imp}_g(a))$$
$$(a, \mathtt{inr}(b)) \mapsto \mathtt{evid}_B(b).$$

The lower half $\zeta \circ \beta$ is also defined by case analysis:

$$(a, \mathtt{inl}(g)) \mapsto \mathtt{imp}_{\lambda x.\,\mathtt{efq}_B(g(x))}(a)$$
$$(a, \mathtt{inr}(b)) \mapsto \mathtt{imp}_{\lambda x.b}(a).$$

Now, to show that $\mathtt{evid}_B \circ \alpha = \zeta \circ \beta$, we must produce an inhabitant of

$$p : \prod_{z:\mathbf{Prob}(A)\times(\neg A+B)} (\mathtt{evid}_B \circ \alpha)(z) = (\zeta \circ \beta)(z).$$

By the elimination rule for product and coproduct types, we can construct such an inhabitant $p$ by considering $z = (a, \mathtt{inl}(g))$ and $z = (a, \mathtt{inr}(b))$ for $a : \mathbf{Prob}(A)$, $g : \neg A$, and $b : B$, and defining

$$p_1 : \prod_{a:\mathbf{Prob}(A)} \prod_{g:\neg A} (\mathtt{evid}_B(\mathtt{efq}_B(\mathtt{imp}_g(a))) = \mathtt{imp}_{\lambda x.\,\mathtt{efq}_B(g(x))}(a))$$
$$p_2 : \prod_{a:\mathbf{Prob}(A)} \prod_{b:B} (\mathtt{imp}_{\lambda x.b}(a) = \mathtt{evid}_B(b)).$$

The conclude $\mathtt{evid}_B \circ \alpha = \zeta \circ \beta$ by the axiom of function extensionality (e.g., Axiom 2.9.3 in [Uni13]). For $p_1$, we appeal to the second elimination rule for the probability type (the rule of inference (7)) to compute

$$\mathtt{imp}_{\lambda x.\,\mathtt{efq}_B(g(x))}(\mathtt{evid}_A(x)) \equiv \mathtt{evid}_B(\mathtt{efq}_B(g(x))), \quad x : A.$$

Now, given $g : \neg A$, $x : A$, and $a : \mathbf{Prob}(A)$, we have

$$q(a, g) \equiv \mathtt{efq}_{\mathtt{imp}_g(a)=g(a)}(\mathtt{imp}_g(a)) : \mathtt{imp}_g(a) = g(x).$$

The proof $p_1$ follows by continuity of functions in type theory:

$$p_1 \equiv \lambda a.\lambda g.\, \mathtt{ap}_{\mathtt{evid}_B \circ \mathtt{efq}_B}(q(a, g)) : \prod_{a:\mathbf{Prob}(A)} \prod_{g:\neg A} (\mathtt{evid}_B(\mathtt{efq}_B(\mathtt{imp}_g(a))) = \mathtt{evid}_B(\mathtt{efq}_B(g(x)))),$$

where $\mathtt{ap}_{\mathtt{evid}_B \circ \mathtt{efq}_B}(q(a, g))$ is the application of $\mathtt{evid}_B \circ \mathtt{efq}_B$ to the path $q(a, g)$, as defined in Lemma 2.2.1 in [Uni13]. For $p_2$, we observe that $\mathtt{imp}_{\lambda x.b}(a) \equiv \mathtt{evid}_B(b)$ so that

$$p_2 :\equiv \lambda a.\lambda b.\mathtt{refl}_{\mathtt{evid}_B(b)}.$$

The conclusion follows by the elimination rules for $\mathbf{Prob}(A) \times (\neg A + B)$ and $\neg A + B$.

To show that the bottom square commutes, we have to prove that $\epsilon \circ \zeta = \delta \circ \gamma$ holds in $\mathbf{Prob}(A) \times (A \to B) \to \mathbf{Prob}(\neg\neg B)$. For the upper half, we have

$$\lambda a.\lambda f.\mathtt{imp}_{\lambda y:B.\lambda g:\neg B.g(y):\mathbf{0}}(\mathtt{imp}_f(a)) : \mathbf{Prob}(A) \times (A \to B) \to \mathbf{Prob}(\neg\neg B),$$

which for $x : A$ and $f : A \to B$ satisfies

$$\begin{aligned}\mathtt{imp}_{\lambda y:B.\lambda g:\neg B.g(y):\mathbf{0}}(\mathtt{imp}_f(\mathtt{evid}_A(x))) &\equiv \mathtt{imp}_{\lambda y:B.\lambda g:\neg B.g(y):\mathbf{0}}(\mathtt{evid}_B(f(x))) \\ &\equiv \mathtt{evid}_{\neg\neg B}(\lambda g : \neg B.g(f(x))).\end{aligned}$$

For the bottom half, we have

$$(\delta \circ \gamma)(a, f) \equiv \mathtt{imp}_{\lambda x.\lambda g.g(f(x))}(a),$$

which for $x : A$ satisfies

$$\mathtt{imp}_{\lambda x.\lambda g.g(f(x))}(\mathtt{evid}_A(x)) \equiv \mathtt{evid}_{\neg\neg B}(\lambda g : \neg\neg B.g(f(x))).$$

The bottom square commutes by reflexivity. The outside square commutes by path concatenation and associativity.

□

*Proof of* (31). Let $p : A \simeq B$ and let $\mathbf{Prob} : \mathcal{U} \to \mathcal{U}$ be the probability type former. By the univalence axiom of HoTT we have $\mathtt{ua}(p) : (A =_{\mathcal{U}} B)$. By [Uni13, Lemma 2.2.1] we have a map

$$\mathtt{ap}_{\mathbf{Prob}} : (A =_{\mathcal{U}} B) \to (\mathbf{Prob}(A) =_{\mathcal{U}} \mathbf{Prob}(B)),$$

which combines with $\mathtt{ua}(p) : A =_{\mathcal{U}} B$ to give

$$\mathtt{ap}_{\mathbf{Prob}}(\mathtt{ua}(p)) : \mathbf{Prob}(A) =_{\mathcal{U}} \mathbf{Prob}(B).$$

Finally, by $\mathtt{idtoequiv} : (\mathbf{Prob}(A) =_{\mathcal{U}} \mathbf{Prob}(B)) \to (\mathbf{Prob}(A) \simeq \mathbf{Prob}(B))$, we obtain

$$\lambda p.\mathtt{idtoequiv}(\mathtt{ap}_{\mathbf{Prob}}(\mathtt{ua}(p))) : (A \simeq B) \to (\mathbf{Prob}(A) \simeq \mathbf{Prob}(B)).$$

□

*Proof of* (33). Recall the definition of $\mathbf{Prob}(B \mid -) : \mathbf{Prob}(A) \to \mathcal{U}$ by

$$\mathbf{Prob}(B \mid a) :\equiv \sum_{y:\mathbf{Prob}(A\times B)} (\mathtt{imp}_{\mathtt{pr}_A}(y) =_{\mathbf{Prob}(A)} a).$$

Now define

$$g \equiv h :\equiv \lambda\langle a, x\rangle.\, \mathtt{pr}_1(x) : \sum_{a:\mathbf{Prob}(A)} \mathbf{Prob}(B \mid a) \to \mathbf{Prob}(A \times B),$$

where $\mathtt{pr}_1 : \mathbf{Prob}(B \mid a) \to \mathbf{Prob}(A \times B)$ is defined as the projection onto the first coordinate of the $\sum$-type $\mathbf{Prob}(B \mid a) \equiv \sum_{y:\mathbf{Prob}(A\times B)}(\mathtt{imp}_{\mathtt{pr}_A}(y) =_{\mathbf{Prob}(A)} a)$,

$$\mathtt{pr}_1 : \sum_{y:\mathbf{Prob}(A\times B)} (\mathtt{imp}_{\mathtt{pr}_A}(y) =_{\mathbf{Prob}(A)} a) \to \mathbf{Prob}(A \times B)$$

$$\mathtt{pr}_1(\langle y, p\rangle) :\equiv y.$$

We construct an inhabitant of

$$(g \circ \mathtt{cond}) \sim \mathtt{id}_{\mathbf{Prob}(A\times B)} :\equiv \prod_{y:\mathbf{Prob}(A\times B)} ((g \circ \mathtt{cond})(y) =_{\mathbf{Prob}(A\times B)} y)$$

by observing that $\mathtt{cond}(y) \equiv \langle \mathtt{imp}_{\mathtt{pr}_A}(y), \langle y, \mathtt{refl}_{\mathtt{imp}_{\mathtt{pr}_A}(y)}\rangle\rangle$, whence $g(\mathtt{cond}(y)) \equiv y : \mathbf{Prob}(A \times B)$ and

$$\lambda y.\mathtt{refl}_y : (g \circ \mathtt{cond}) \sim \mathtt{id}_{\mathbf{Prob}(A\times B)}\,.$$

It remains to prove that

$$(\mathtt{cond} \circ h) \sim \mathtt{id}_{\sum_{a:\mathbf{Prob}(A)} \mathbf{Prob}(B|a)}\,.$$

Note first that

$$(\mathtt{cond} \circ h)(\langle a, \langle y, p\rangle\rangle) :\equiv \langle \mathtt{imp}_{\mathtt{pr}_A}(y), \langle y, \mathtt{refl}_{\mathtt{imp}_{\mathtt{pr}_A}(y)}\rangle\rangle.$$

By the elimination rule for $\sum$-types, it is enough to prove

$$\prod_{\langle a,\langle y,p\rangle\rangle:\sum_{a:\mathbf{Prob}(A)}\mathbf{Prob}(B|a)} \langle a, \langle y, p\rangle\rangle = \langle \mathtt{imp}_{\mathtt{pr}_A}(y), \langle y, \mathtt{refl}_{\mathtt{imp}_{\mathtt{pr}_A}(y)}\rangle\rangle.$$

First note that

$$\sum_{a:\mathbf{Prob}(A)} \mathbf{Prob}(B \mid a) \simeq \sum_{z:\mathbf{Prob}(A)\times\mathbf{Prob}(A\times B)} (\mathtt{imp}_{\mathtt{pr}_A}(\mathtt{pr}_{\mathbf{Prob}(A\times B)}(z)) = \mathtt{pr}_A(z)).$$

By the elimination rule for product types, we can assume $z = \langle a, y\rangle$ for $a : \mathbf{Prob}(A)$ and $y : \mathbf{Prob}(A \times B)$ so that

$$\sum_{a:\mathbf{Prob}(A)} \mathbf{Prob}(B \mid a) \simeq \sum_{\langle a,y\rangle:\mathbf{Prob}(A)\times\mathbf{Prob}(A\times B)} (\mathtt{imp}_{\mathtt{pr}_A}(y) = a).$$

Now, given any $\langle a, \langle y, p\rangle\rangle : \sum_{a:\mathbf{Prob}(A)} \mathbf{Prob}(B \mid a)$, we immediately have $p : \mathtt{imp}_{\mathtt{pr}_A}(y) = a$, and thus $p^{-1} : a = \mathtt{imp}_{\mathtt{pr}_A}(y)$, $\mathtt{refl}_y : y = y$, and

$$\mathrm{pair}^{=}(p^{-1}, \mathtt{refl}_y) : \langle a, y\rangle = \langle \mathtt{imp}_{\mathtt{pr}_A}(y), y\rangle,$$

with $\mathtt{pair}^{=}$ as defined in [Uni13, Theorem 2.6.2]. By Theorem 2.7.2 in [Uni13], it remains to show that

$$\mathtt{transport}^C(\mathtt{pair}^{=}(p^{-1}, \mathtt{refl}_y), p^{-1}) = \mathtt{refl}_{\mathtt{imp}_{\mathtt{pr}_A}(y)}, \tag{45}$$

for $C : \mathbf{Prob}(A) \times \mathbf{Prob}(A \times B) \to \mathcal{U}$ defined by

$$C(\langle a, y\rangle) :\equiv a =_{\mathbf{Prob}(A)} \mathtt{imp}_{\mathtt{pr}_A}(y).$$

We argue by based path induction as follows.

Fix $\langle a, \langle y, p\rangle\rangle : \sum_{a:\mathbf{Prob}(A)} \mathbf{Prob}(B \mid a)$ and define

$$D : \prod_{\langle a',y'\rangle:\mathbf{Prob}(A)\times\mathbf{Prob}(A\times B)} (\langle a', y'\rangle = \langle \mathtt{imp}_{\mathtt{pr}_A}(y), y\rangle) \to \mathcal{U}$$

$$D(\langle a', y'\rangle, p') :\equiv \mathtt{transport}^C(p', \mathtt{ap}_{\mathtt{pr}_A}(p')) = \mathtt{refl}_{\mathtt{imp}_{\mathtt{pr}_A}(y)}.$$

Arguing by based path induction at $\langle \mathtt{imp}_{\mathtt{pr}_A}(y), y\rangle$, we can assume $\langle a', y'\rangle \equiv \langle \mathtt{imp}_{\mathtt{pr}_A}(y), y\rangle$, so that

$$\begin{aligned} &D(\langle \mathtt{imp}_{\mathtt{pr}_A}(y), y\rangle, \mathtt{refl}_{\langle \mathtt{imp}_{\mathtt{pr}_A}(y), y\rangle}) \equiv \\ &\quad \equiv \mathtt{transport}^C(\mathtt{refl}_{\langle \mathtt{imp}_{\mathtt{pr}_A}(y), y\rangle}, \mathtt{ap}_{\mathtt{pr}_A}(\mathtt{refl}_{\langle \mathtt{imp}_{\mathtt{pr}_A}(y), y\rangle})) = \mathtt{refl}_{\mathtt{imp}_{\mathtt{pr}_A}(y)} \\ &\quad \equiv \mathtt{ap}_{\mathtt{pr}_A}(\mathtt{refl}_{\langle \mathtt{imp}_{\mathtt{pr}_A}(y), y\rangle}) = \mathtt{refl}_{\mathtt{imp}_{\mathtt{pr}_A}(y)}, \end{aligned}$$

for which we have an inhabitant by the propositional computation rule for $\mathtt{pair}^{=}$; see [Uni13, p. 106].

By based path induction, we have an inhabitant of $D(z, p')$ for every $z : \mathbf{Prob}(A) \times \mathbf{Prob}(A \times B)$ and $p' : z = \langle \mathtt{imp}_{\mathtt{pr}_A}(y), y\rangle$. In particular, for $z \equiv \langle a, y\rangle$, with $a :$

$\mathbf{Prob}(A)$, $y : \mathbf{Prob}(A \times B)$, and $p' \equiv \mathtt{pair}^{=}(p^{-1}, \mathtt{refl}_y)$, we have an inhabitant

$$d : D(\langle a, y\rangle, \mathtt{pair}^{=}(p^{-1}, \mathtt{refl}_y)) \equiv$$
$$\equiv \mathtt{transport}^C(\mathtt{pair}^{=}(p^{-1}, \mathtt{refl}_y), \mathtt{ap}_{\mathtt{pr}_A}(\mathtt{pair}^{=}(p^{-1}, \mathtt{refl}_y))) = \mathtt{refl}_{\mathtt{imp}_{\mathtt{pr}_A}(y)}.$$

Again by the propositional computation rule for $\mathtt{pair}^{=}$, we have an inhabitant

$$r : \mathtt{ap}_{\mathtt{pr}_A}(\mathtt{pair}^{=}(p^{-1}, \mathtt{refl}_y)) = p^{-1}.$$

And thus, by applying the transport function to the path $r$, we have

$$\mathtt{ap}_{\mathtt{transport}^C(\mathtt{pair}^{=}(p^{-1}, \mathtt{refl}_y), -)}(r) :$$
$$\mathtt{transport}^C(\mathtt{pair}^{=}(p^{-1}, \mathtt{refl}_y), \mathtt{ap}_{\mathtt{pr}_A}(\mathtt{pair}^{=}(p^{-1}, \mathtt{refl}_y))) =$$
$$= \mathtt{transport}^C(\mathtt{pair}^{=}(p^{-1}, \mathtt{refl}_y), p^{-1}).$$

By path concatenation, we obtain

$$\mathtt{ap}_{\mathtt{transport}^C(\mathtt{pair}^{=}(p^{-1}, \mathtt{refl}_y), -)}(r)^{-1} \bullet d :$$
$$\mathtt{transport}^C(\mathtt{pair}^{=}(p^{-1}, \mathtt{refl}_y), p^{-1}) = \mathtt{refl}_{\mathtt{imp}_{\mathtt{pr}_A}(y)},$$

as required by (45).

□

## References

[AGM85] Carlos Alchourron, Peter Gärdenfors, and David Makinson, *On the logic of theory change: Partial meet contraction and revision functions*, The Journal of Symbolic Logic **50** (1985), no. 2, 510–530.

[APW13] S. Awodey, A. Pelayo, and M.A. Warren, *Voevodsky's Univalence Axiom in Homotopy Type Theory*, Notices of the AMS **60** (2013), no. 9, 1164–1167.

[Awo14] S. Awodey, *Structuralism, Invariance, and Univalence*, Philosophia Mathematica **22** (2014), no. 1, 1–11.

[Awo16] ———, *Univalence as a Principle of Logic*, https: (2016).

[Bro81] L.E.J. Brouwer, *Brouwer's cambridge lectures on intuitionism*, Cambridge University Press, 1981.

[Broa] ———.

[Brob] ———, *Intuitionism and Formalism*, Bulletin (New Series) of the American Mathematical Society **37**, no. 1, 55–64.

[Car50] Rudolph Carnap, *Logical Foundations of Probability*, University of Chicago Press, 1950.

[CF] H. Curry and R. Feys, Craig, william, ed., combinatory logic vol. i.

[Cra18] Harry Crane, *Logic of Probability and Conjecture (full version)*, in preparation (2018).

[Cur34] H. Curry, *Functionality in Combinatory Logic*, Proceedings of the National Academy of Sciences **20** (1934), 584–590.

[Dem67] Arthur P. Dempster, *Upper and lower probabilities induced by a multivalued mapping*, The Annals of Mathematical Statistics **38** (1967), no. 2, 325–339.

[dF] B. de Finetti, *La prévision: ses lois logiques, ses sources subjectives*, Annales de l'Institut Henri Poincaré **7**, 1–68.

[Dum00] M. Dummett, *Elements of intuitionism, 2nd edition*, Oxford Logic Guides (Book 39), Clarendon Press, 2000.

[Fol92] Richard Foley, *Working Without a Net*, Oxford University Press, 1992.

[Fra15] James Franklin, *The Science of Conjecture: Evidence and Probability before Pascal*, Johns Hopkins University Press, Baltimore, MD, USA, 2015.

[GZP+89] Gerd Gigerenzer, Zeno Zwijtink, Theodore Porter, Lorraine Daston, John Beatty, and Lorenz Kruger, *The Empire of Chance: How Probability Changed Science and Everyday Life (Ideas in Context)*, Cambridge University Press, 1989.

[Hac75] I. Hacking, *The Emergence of Probability: A Philosophical Study of Early Ideas about Probability, Induction and Statistical Inference, 2nd edition*, Cambridge Series on Statistical & Probabilistic Mathematics, Cambridge University Press, 1975.

[Hey71] A. Heyting, *Intuitionism: An introduction*, Study in Logic & Mathematics, 1971.

[How69] W.A. Howard, *The formulae-as-types notion of construction*, in seldin, jonathan p.; hindley, j. roger, to h.b. curry: Essays on combinatory logic, lambda calculus and formalism, 1969, pp. 479–490.

[KL11] Chris Kapulkin and Peter LeFanu Lumsdaine, *The Simplicial Model of Univalent Foundations (after Voevodsky)*, arXiv:1211.2851 (2011).

[LP14] J. Ladyman and S. Presnell, *A Primer on Homotopy Type Theory Part 1: The Formal Type Theory*, http: (2014).

[Maz12] Barry Mazur, *Is it plausible?*, Available at http://www.math.harvard.edu/~mazur/papers/Plausibility.Notes.3.pdf (2012).

[ML84] P. Martin-Löf, *Intuitionistic type theory*, Studies in Proof Theory. Lecture Notes, v. 1, 1984.

[ML87] ______, *Truth of a Proposition, Evidence of a Judgement, Validity of a Proof*, Synthese **73** (1987), 407–420.

[ML96] ______, *On the Meanings of the Logical Constants and the Justifications of the Logical Laws*, Nordic Journal of Philosophical Logic **1** (1996), no. 1, 11–60.

[Pól54] George Pólya, *Mathematics and Plausible Reasoning*, Martino Publishing, Mansfield Centre, CT, 1954.

[Por96] Theodore Porter, *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*, Princeton University Press, 1996.

[PW12] Á. Pelayo and M.A. Warren, *Homotopy Type Theory and Voevodsky's Univalent Foundations*, arXiv1210:5658 (2012).

[Ram26] Frank P. Ramsey, *Truth and Probability* (1926).

[SF18] Ted Shear and Branden Fitelson, *Two approaches to belief revision*, Erkenntnis **in press** (2018).

[Sha76] Glenn R. Shafer, *A mathematical theory of evidence*, Princeton University Press, 1976.

[Shu17] M. Shulman, *Homotopy type theory: the logic of space*, arXiv:1703.03007 (2017).

[Tse16] Dimitis Tsementzis, *Univalent Foundations as Structuralist Foundations*, Synthese (2016).

[Tse17] Dimitris Tsementzis, *A meaning explanation for hott*, http: (2017).
https://homotopytypetheory.org/book

[Uni13] Univalent Foundations Program, The, *Homotopy type theory: Univalent foundations of mathematics*, https://homotopytypetheory.org/book, Institute for Advanced Study, 2013.

[Wea03] Brian Weatherson, *From Classical to Intuitionistic Probability*, Notre Dame Journal of Formal Logic **44** (2003), 111–123.

[Wit73] Ludwig Wittgenstein, *Philosophical investigations (3rd edition)*, Pearson, 1973.



Department of Statistics & Biostatistics, Rutgers University, 110 Frelinghuysen Road, Piscataway, NJ 08854, USA

*E-mail address*: hcrane@stat.rutgers.edu