A Formal Model for Intuitive Probabilistic Reasoning

Harry Crane

Department of Statistics
Rutgers

March 25, 2019

Motivation

1. Some common statements of belief:
   - **Law**: I believe O.J. Simpson is not guilty of murder (beyond reasonable doubt) because the glove didn't fit.
   - **Mathematics**: I believe Goldbach's conjecture is probably true because it has been verified for $> 4 \times 10^{18}$ special cases.[1]
   - **Science**: There will be a partial solar eclipse on June 24, 2112 because that's what the laws of physics and relevant theories of planetary motion predict.[2]
   - **Common Sense**: It's safe to cross the street because there is no car within 200 yards.
2. These statements
   - rely on intuition about when it is reasonable to believe something,
   - involve probabilistic judgment, i.e., a judgment about what is 'probably true' in light of evidence, and
   - give reasons to justify the main claim.
3. All of these statements fall under heading of *intuitive probabilistic reasoning* (IPR).

[1] Goldbach's conjecture: every even integer greater than 3 is the sum of 2 primes, e.g., $3 + 1 = 4$, $3 + 3 = 6$, $5 + 3 = 8$, ...

[2] https://eclipse.gsfc.nasa.gov/SEcat5/SE2101-2200.html

Logic of Justified Belief

- All of the previous statements have the form '*A* because *a*' where

  A is a claim (something I believe).

  a is a reason (justification) for the claim.

- These statements
  - convey subjective beliefs but also provide an external qualification of that belief;
  - do not have the form of Bayesian credences.

Objectives:

- Most logical reasoning involves conjecture, intuition, and partial information, but not many facts. IPR seeks to formalize this process.
- IPR is not concerned with whether a given judgment is 'true', 'correct', etc.
- IPR concerns the logic by which such judgments are made, and what it is that makes intuitive judgments 'sound'.

Justified Beliefs about Uncertain Claims

- Justified beliefs correspond to a judgment/claim that something is
  - True
  - Probable
  - Plausible
  - Typical
  - Likely
  - ...
- Need not be quantified. Often cannot be.
- I don't consider whether claims can be justified in an objective sense, but rather the way in which claims are justified subjectively.
- Justification doesn't involve multiple 'possible worlds'. All claims are made about **this** world and the justification pertains to **this** world. ('Possible worlds' reasoning may be one of the many ways that claims can be justified.)

Examples:

- Relevant domains include: law, mathematical practice, scientific practice, common sense.
- Topical example: Use of '$P < 0.05$' to justify conclusions in scientific work. What makes these conclusions justified?

Quantifying uncertainty, Qualifying belief

Quantifying uncertainty (Bayesian)

- When uncertain about a decidable statement (e.g., 'it will rain tomorrow'), I can quantify that uncertainty as 'degrees of belief', 'credences', etc. and update/calculate using the probability calculus.
- Approach taken in statistics, machine learning, scientific analysis, etc.
- Quantification of undecidable statements can be hard to interpret. Or, they may be too easy to 'interpret', e.g., *P*-values and replication crisis.

Qualifying belief (IPR)

- When making a claim about something undecidable, subjective, counterfactual, or unquantifiable, I can qualify my claim by giving a reason for it.
- Approach taken in law, mathematical conjecture, scientific practice, and everyday decision making.

These two settings are consistent with one another. Will discuss later.

Why not Bayes all the way?

Short Answer: Because it's not enough.
Longer Answer:

- The probability calculus expresses beliefs numerically by a probability function.

  For a statement A, $P(A)$ is an agent's degree of belief in A.

- Operational interpretation: $P(A)$ determines the odds at which an agent would be indifferent between a bet for or against A.
- Fine in some cases. Not all (or most).
- Many claims we deal with may not be decidable.
  - There may be* a fact of the matter whether O.J. is guilty, but we can never know the truth.
  - There may be* a fact of the matter whether Goldbach's conjecture is true, but we may never find out.
  - There may be* a fact of the matter whether there will be an eclipse in June 2112, but we won't live long enough to see it.
- Assigning a credence to these outcomes is difficult/impossible and more importantly (i) unhelpful for communicating with other agents and (ii) not the way real people express their beliefs.
- Bayes is unhelpful because if you assign credence $P(A) = 0.95$, I can be pretty sure that you believe A but I have no idea why you believe it. So I don't know how to update my beliefs based on yours.
- Real people communicate by giving reasons for their beliefs.

Evidence is primitive

1. Probability raising: E is evidence for H if $P(H \mid E) > P(H)$.

   $\Rightarrow$ Beliefs first, evidence second.

2. IPR: Evidence is primitive.

   $\Rightarrow$ Evidence is the basis of belief.

3. In the examples:
   1. **Law**: The trial (evidence) comes before the verdict (judgment).
   2. **Math**: Empirical evidence (as in Goldbach's conjecture) or other intuitive rationale comes before the conjecture.
   3. **Science**: Empirical evidence often comes before the hypothesis. Experimentation comes before the conclusion.

Formalism for IPR (Overview)

Formal:

- Syntax: Martin-Löf type theory (MLTT)
- Semantics: Homotopy type theory (HoTT)
- Extra structure: A type former **Bel** on top of usual rules of MLTT

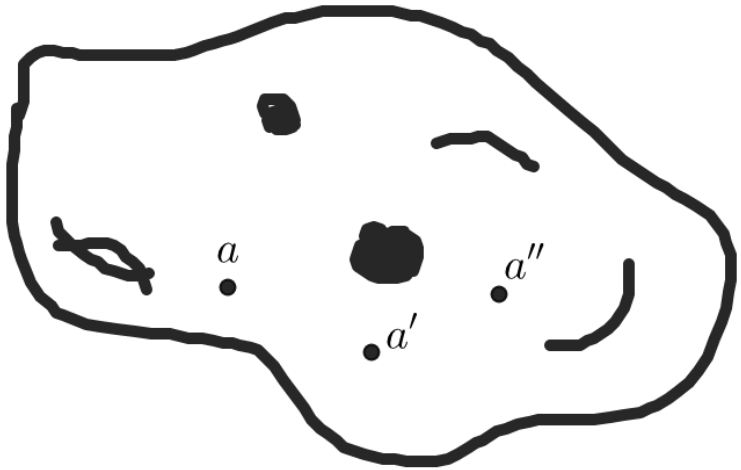
Pre-formal:

- Syntax: Rules for expressing judgments of the form '*A* because *a*'.
- Semantics: Subjective judgments reflect agent's credal state (frame of mind, context).
- Extra structure: Expressions about uncertain claims ('Probably *A* because *a*').

Basic Objects

Two basic objects:

Notation	Object (MLTT)	HoTT	IPR
A	type	homotopy type	claim
a	term	point	proof/evidence



A : **Type** Space representing a claim

a, a', a'' : A Points of A representing pieces of evidence or proofs.

Basic judgments

Judgment	MLTT	IPR
A : **Type**	A is a type	A is a claim
$a : A$	a is a term of A	a is evidence for A
$A \equiv A'$: **Type**	A and A' are identical types	A and A' are identical claims
$a \equiv a' : A$	a and a' are identical terms of A	a, a' are identical evidence for A

Technical point:

- The latter two judgments ($A \equiv A'$: **Type** and $a \equiv a' : A$) are *judgmental identities*. A and A', a and a' refer to the same thing (literally).
- There is another notion of *propositional identity* in MLTT, where two non-identical objects are claimed to be identical. (Won't discuss this much today.)
- Example:

It's raining now.

现在下雨

Contexts and Judgments

Contexts:

Judgment	MLTT	IPR
Δ **ctx**	Δ is a context	Δ is a state of mind, frame of reference

Role of context:

- All judgments are of the form above, and are made relative to a context:

$$\begin{array}{ccc} \Delta & \vdash & \mathcal{J} \\ \text{Context} & \vdash & \text{Judgment} \end{array}$$

- Formal rules for contexts imply: Every well-formed context Δ is a list (possibly empty) of primitive judgments:

$$\Delta \equiv (a_1 : A_1, \ldots, a_n : A_n).$$

- Interpretation: Frame of reference for subsequent judgments is the accumulation of prior judgments.

Propositions-as-types (Curry–Howard)

- Numerous semantics for objects of MLTT

	terms	types	$a : A$
Set theory	element	set	$a \in A$
Topology/HoTT	point	space	a is a point of A
BHK	algorithm/solution	problem	a is an algorithm/solution to A
Curry–Howard	proof	proposition	a is a proof of A
IPR	reason	claim	A because a

- IPR expands on Curry–Howard to include generic (non-mathematical) claims.
- Add a belief type **Bel** on top of other rules of MLTT.

Rules of MLTT

Syntax of MLTT determined by rules of

1. Formation: A rule for forming a new type out of old terms and types.
2. Introduction: Rule(s) for introducing terms of the new type.
3. Elimination: Rule(s) for defining functions from the new type.
4. Computation: Rule(s) for computing terms derived from introduction and elimination rules.

Orientation:

- In MLTT, there are no “axioms”.
- Rules determine what logical operations are permitted given certain inputs.
- UF/HoTT = MLTT + Univalence Axiom (UA).

Types in MLTT

MLTT	PL/FOL	Set Theory	Interpretation
$A \times B$	$A \wedge B$	$A \cap B$	A and B
$A + B$	$A \vee B$	$A \cup B$	A or B
$\mathbf{0}$	$\bot$	$\emptyset$	contradiction
$A \to \mathbf{0}$	$\neg A$	A^c	not A
$A \to B$	$\neg A \vee B$	$A \subseteq B$	if A then B
$\prod_{a:A} B(a)$	$\forall a B(a)$	$\bigcap_{a \in A} B(a)$	$B(a)$ for all $a : A$
$\sum_{a:A} B(a)$	$\exists a B(a)$	$\bigcup_{a \in A} B(a)$	there exists a such that $B(a)$

- Note: Correspondence between MLTT, PL, and set theory is not exact. Just a rough guideline.
- To understand the difference need to look at the rules of formation, introduction, elimination, and computation in MLTT.

Product Type ($\times$)

- **Formation**:

$$\frac{\Delta \vdash A : \textbf{Type} \qquad \Delta \vdash B : \textbf{Type}}{\Delta \vdash A \times B : \textbf{Type}} \quad (\times\text{-form})$$

- **Introduction**:

$$\frac{\Delta \vdash A : \textbf{Type} \qquad \Delta \vdash B : \textbf{Type}}{\Delta, a : A, b : B \vdash \mathtt{pair}(a, b) : A \times B} \quad (\times\text{-intro})$$

- **Elimination**:

$$\frac{\begin{array}{c}\Delta, x : A \times B \vdash C : \textbf{Type} \\ \Delta, a : A, b : B \vdash d(a, b) : C\end{array}}{\Delta, x : A \times B \vdash \mathtt{split}_d(x) : C} \quad (\times\text{-elim})$$

- **Computation**:

$$\frac{\begin{array}{c}\Delta, x : A \times B \vdash C : \textbf{Type} \\ \Delta, a : A, b : B \vdash d(a, b) : C\end{array}}{\Delta, a : A, b : B \vdash \mathtt{split}_d(\mathtt{pair}(a, b)) \equiv d(a, b) : C} \quad (\times\text{-comp})$$

- Rules determine the meaning of '$A \times B$' implicitly.

The **Bel** Type

(1) **Formation***: If one can claim to know A, then one can claim to believe A:*

$$\frac{\Delta \vdash A : \mathbf{Type}}{\Delta \vdash \mathbf{Bel}(A) : \mathbf{Type}} \quad \mathbf{Bel}\textit{-form}$$

(2) **Introduction***: If A has been proven, then one is justified in believing A:*

$$\frac{\Delta \vdash A : \mathbf{Type}}{\Delta, a : A \vdash \mathtt{evid}_A(a) : \mathbf{Bel}(A)} \quad \mathbf{Bel}\textit{-intro}$$

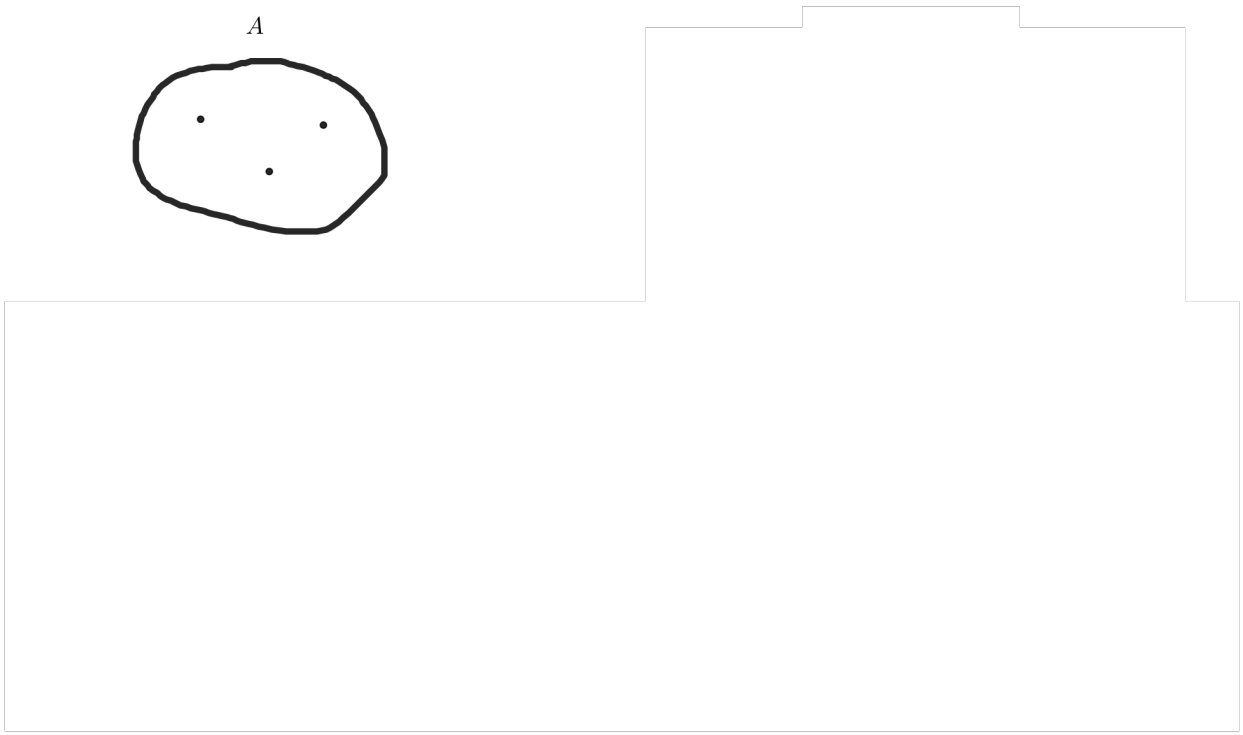
(3) **Elimination***: If A implies B, then justified belief in A justifies belief in B:*

$$\frac{\Delta \vdash A : \mathbf{Type} \qquad \Delta \vdash B : \mathbf{Type} \qquad \Delta, a : A \vdash d(a) : B}{\Delta, x : \mathbf{Bel}(A) \vdash \mathtt{imp}_d(x) : \mathbf{Bel}(B)} \quad \mathbf{Bel}\textit{-elim}$$

(4) **Computation***: Operations (2) and (3) commute:*

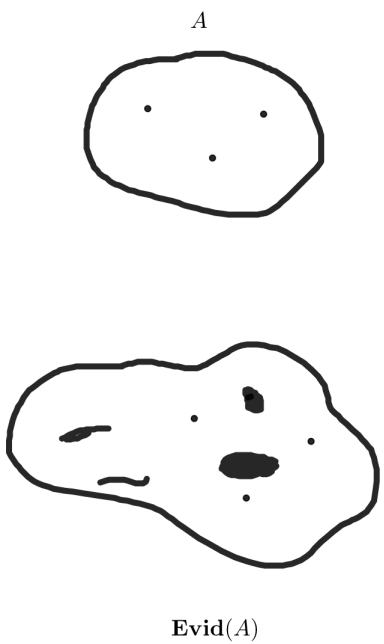
$$\frac{\Delta \vdash A : \mathbf{Type} \qquad \Delta \vdash B : \mathbf{Type} \qquad \Delta, a : A \vdash d(a) : B}{\Delta, a : A \vdash \mathtt{imp}_d(\mathtt{evid}_A(a)) \equiv \mathtt{evid}_B(d(a)) : \mathbf{Bel}(B)} \quad \mathbf{Bel}\textit{-comp}$$

$$\frac{A : \textbf{Claim}}{} \quad (1)$$



* **Evid**$(\bullet) \equiv$ **Bel**$(\bullet)$ and **Claim** $\equiv$ **Type** in above picture.

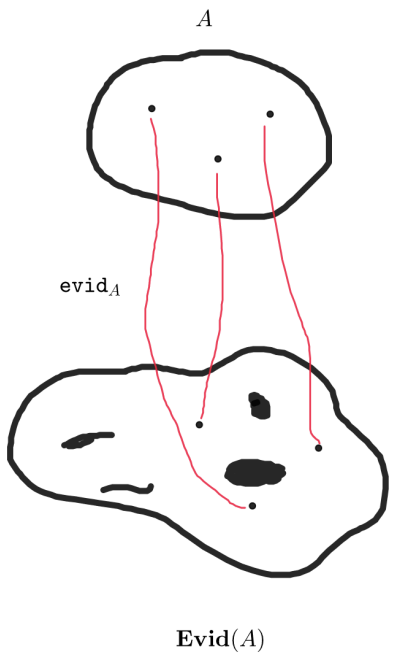
$$\frac{A : \textbf{Claim}}{\textbf{Bel}(A) : \textbf{Claim}} \quad (1)$$



* **Evid**$(\bullet) \equiv$ **Bel**$(\bullet)$ and **Claim** $\equiv$ **Type** in above picture.

The Logic of Evidence

$$\frac{A : \textbf{Claim}}{\textbf{Bel}(A) : \textbf{Claim}} \quad (1) \qquad \frac{a : A}{\texttt{evid}_A(a) : \textbf{Bel}(A)} \quad (2)$$

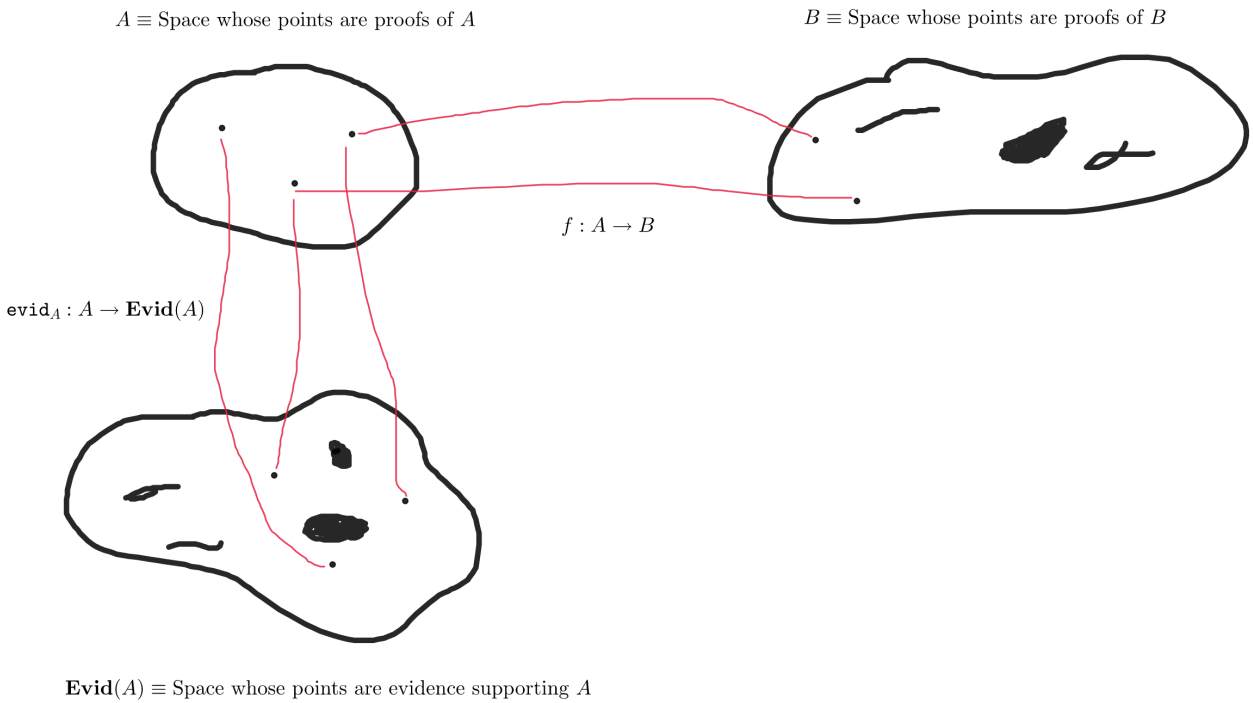


* $\textbf{Evid}(\bullet) \equiv \textbf{Bel}(\bullet)$ and $\textbf{Claim} \equiv \textbf{Type}$ in above picture.

The Logic of Evidence

$$\frac{A : \textbf{Claim}}{\textbf{Bel}(A) : \textbf{Claim}} \quad (1) \qquad \frac{a : A}{\texttt{evid}_A(a) : \textbf{Bel}(A)} \quad (2)$$

$$\frac{A : \textbf{Claim} \qquad B : \textbf{Claim}}{f : A \to B} \quad (3)$$

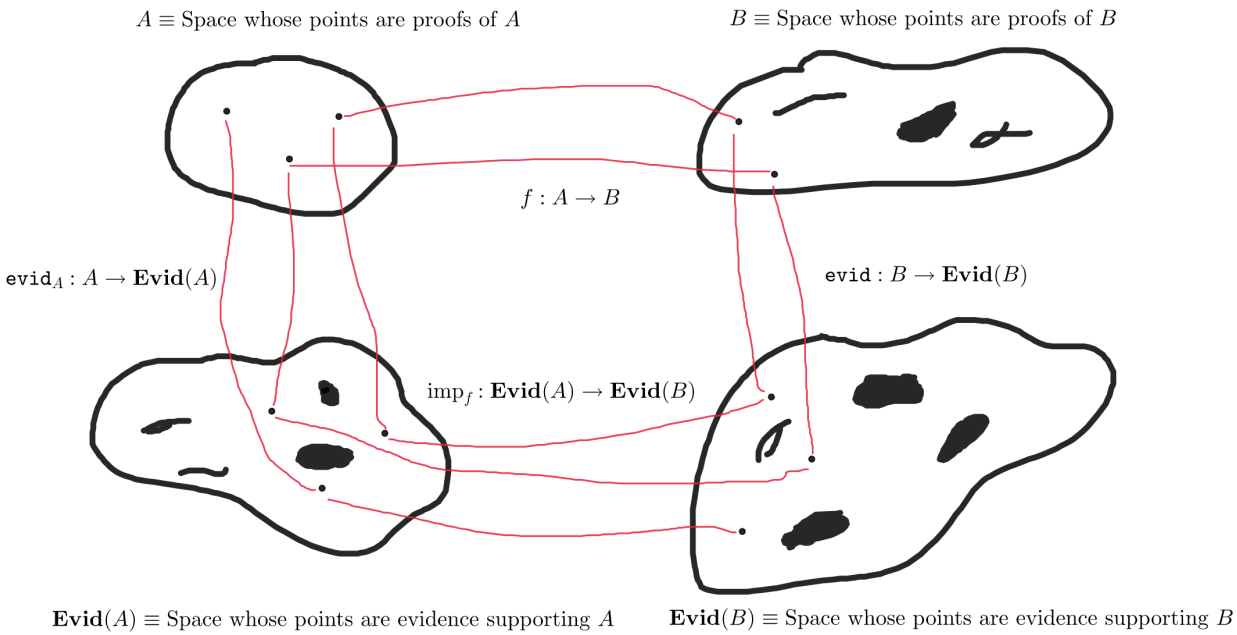


* **Evid**$(\bullet) \equiv$ **Bel**$(\bullet)$ and **Claim** $\equiv$ **Type** in above picture.

The Logic of Evidence

$$\frac{A : \textbf{Claim}}{\textbf{Bel}(A) : \textbf{Claim}} \quad (1) \qquad \frac{a : A}{\texttt{evid}_A(a) : \textbf{Bel}(A)} \quad (2)$$

$$\frac{A : \textbf{Claim} \qquad B : \textbf{Claim} \qquad f : A \to B}{\texttt{imp}_f : \textbf{Bel}(A) \to \textbf{Bel}(B)} \quad (3)$$



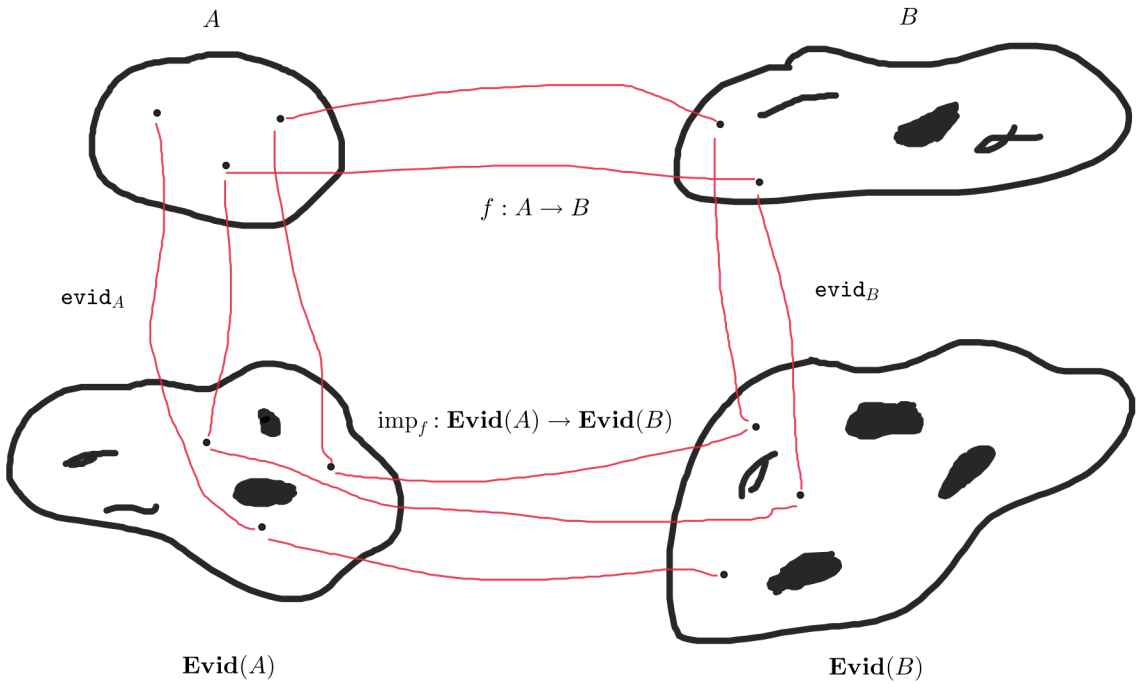
* **Evid**$(\bullet) \equiv$ **Bel**$(\bullet)$ and **Claim** $\equiv$ **Type** in above picture.

The Logic of Evidence: Full Picture

$$\frac{A : \textbf{Claim}}{\textbf{Bel}(A) : \textbf{Claim}} \quad (1) \qquad \frac{a : A}{\texttt{evid}_A(a) : \textbf{Bel}(A)} \quad (2)$$

$$\frac{A : \textbf{Claim} \qquad B : \textbf{Claim} \qquad f : A \to B}{\texttt{imp}_f : \textbf{Bel}(A) \to \textbf{Bel}(B)} \quad (3)$$

$$\frac{A : \textbf{Claim} \qquad B : \textbf{Claim} \qquad f : A \to B}{\texttt{imp}_f \circ \texttt{evid}_A \equiv \texttt{evid}_B \circ f : \textbf{Bel}(A) \to \textbf{Bel}(B)} \quad (4)$$



* **Evid**$(\bullet) \equiv$ **Bel**$(\bullet)$ and **Claim** $\equiv$ **Type** in above picture.

Consequences of these axioms

The following can be derived as **theorems** from the rules (1)-(4).

- $\mathbf{Bel}(A \times B) \to \mathbf{Bel}(A) \times \mathbf{Bel}(B) \to \mathbf{Bel}(A) + \mathbf{Bel}(B) \to \mathbf{Bel}(A + B)$

  Justified belief in 'A and B' (jointly) justifies belief in A and B (individually).
  Justified belief in A and in B justifies belief in A or in B.
  Justified belief in A or in B justifies belief in 'A or B'.
  (Arrows do not reverse in general.)

Universal/Existential quantification:

- $\mathbf{Bel}\ (\Pi_{a:A} B(a)) \to \Pi_{a:A}\ \mathbf{Bel}(B(a))$.

  If belief that $B(a)$ holds for all $a : A$ is justified then for all $a : A$ belief in $B(a)$ in justified.

- $\Sigma_{a:A}\ \mathbf{Bel}(B(a)) \to \mathbf{Bel}\ (\Sigma_{a:A}\ B(a))$.

  If belief in $B(a)$ is justified for some $a : A$ then belief that B(a) holds for some $a : A$ is justified.

Derived inference rules:

- $\mathbf{Bel}(A) \times (A \to B) \to \mathbf{Bel}(A \times B)$

  If belief in A is justified and A implies B then belief in A and B is justified.

- $A \times \mathbf{Bel}(A \to B) \to \mathbf{Bel}(A \times B)$.

Formalization in Coq/UniMath: Definition of **Bel** as a homotopy type

$$\frac{A : \mathbf{Type}}{\mathbf{Bel}(A) : \mathbf{Type}} \quad (1)$$

$$\frac{\begin{array}{c} A : \mathbf{Type} \\ a : A \end{array}}{\mathtt{evid}_A(a) : \mathbf{Bel}(A)} \quad (2)$$

$$\frac{\begin{array}{cc} A : \mathbf{Type} & B : \mathbf{Type} \\ \multicolumn{2}{c}{f : A \to B} \end{array}}{\mathtt{imp}_f : \mathbf{Bel}(A) \to \mathbf{Bel}(B)} \quad (3)$$

$$\frac{\begin{array}{cc} A : \mathbf{Type} & B : \mathbf{Type} \\ \multicolumn{2}{c}{f : A \to B} \end{array}}{\begin{array}{c} \mathtt{imp}_f \circ \mathtt{evid}_A \equiv \mathtt{evid}_B \circ f : \\ : \mathbf{Bel}(A) \to \mathbf{Bel}(B) \end{array}} \quad (4)$$

Written in Coq using UniMath library:

```
(** * Foundations of Data Science
Harry Crane
Started on May 1, 2018
Preparation for BFF 5 talk at University of Michigan on "The Fundamental Principle of Data Science"

Implementation of basic theorems from H. Crane.  (2018).  Logic of probability and conjecture.

https://philpapers.org/rec/CRALOP
*)

(** *** Settings *)

Unset Automatic Introduction.
(* The above line has to be removed for the file to compile with Coq8.2 *)

(** *** Imports *)

Require Export UniMath.Foundations.Preamble.
Require Export UniMath.Foundations.PartA.
Require Export UniMath.Foundations.UnivalenceAxiom.

(* end of "Preamble" *)

(* Define Evidence type former *)

Definition Space := UU.

Inductive Evid (A : Space) : Space :=
| evid : A → Evid A.

(* Elimination, Computation, and Zero rules for Evidence Type *)

Hypothesis  imp: ∏  (A : Space) (P : Evid A → Space),
    (∏ a : A, P (evid A a)) → ∏ p : Evid A, Evid (P p).

Hypothesis  comp : ∏  (A : Space) (P : Evid A → Space)
  (d : (∏ a : A, P (evid A a))), (∏ a:A, (imp A P d) (evid A a) = evid (P (evid A a)) (d a)).

Hypothesis evid_zero : Evid ∅ = ∅.

(* Special case of imp for nondependent types *)

Definition  Prob_imp_nondep : ∏  {A B : Space}, (A → B) → (Evid A → Evid B).
Proof.
  intros A B f a.
  apply (imp A (λ _, B) (λ a, f a)) in a.
  apply a.
Defined.

-U:**-  FPDS-BFF5-1    Top L11    (Coq Script(0-) Holes)
```

Formalization in Coq/UniMath: Proofs of basic properties

Theorems mentioned earlier:

$$\mathbf{Bel}(A \times B) \to \mathbf{Bel}(A) \times \mathbf{Bel}(B)$$

$$\mathbf{Bel}(A) \to \mathbf{Bel}(A + B)$$

$$\mathbf{Bel}(B) \to \mathbf{Bel}(A + B)$$

$$\mathbf{Bel}(A) + \mathbf{Bel}(B) \to \mathbf{Bel}(A + B)$$

Written in Coq using UniMath library:

```
(* Theorems for  product and coproduct *)

Theorem split_prob: ∏  (A B : Space), Evid ( A × B ) → Evid A × Evid B.

Proof. intros A B.  intros x.
       assert (H: Evid (A × B) → Evid A). { intros.
                                            apply (@Prob_imp_nondep (A × B) A pr1) in x. apply x. }

       assert (H': Evid (A × B) → Evid B). {intros.  apply (@Prob_imp_nondep (A × B) B pr2) in x. apply x. }
       split.
       - apply H in x. apply x.
       - apply H' in x. apply x.
Qed.

Notation "A ⨿ B" := (coprod A B) (at level 20): type_scope.

Theorem inl_prob : ∏  {A B : Space}, Evid A → Evid ( A ⨿ B ).

Proof.
  intros A B.  intros x.
  Check Prob_imp_nondep.
  apply (@Prob_imp_nondep A (A ⨿ B) (@inl A B)) in x.
  apply x.
Qed.

Theorem inr_prob : ∏  {A B : Space}, Evid B → Evid ( A ⨿ B ).

Proof.
  intros A B.  intros x.
  apply (Prob_imp_nondep (@inr A B)) in x.
  apply x.
Qed.

Theorem prod_co_prob : ∏  (A B : Space), Evid ( A × B ) → Evid ( A ⨿ B ).

Proof. intros A B.  intros x.
       apply (@Prob_imp_nondep ( A × B ) (coprod A B) (fun y => @inl A B (pr1 y))) in x. apply x.
Qed.

Theorem join_prob : ∏  (A B : Space), Evid A ⨿ Evid B → Evid ( A ⨿ B ).
Proof.
  intros A B. intros x.
  destruct x.
  - apply inl_prob. apply e.
  - apply inr_prob. apply e.
Qed.
```

Lockean Semantics for IPR

- Agent with beliefs given by $P(\cdot)$.
- $1/2 < t \leq 1$ (Lockean threshold)
- $\mathcal{B}$ is the agent's belief set.

(Lockean Thesis)

If $P(A) \geq t$, then $A \in \mathcal{B}$.

- Fix Ω and write $\mathcal{P}(\Omega) = \{\text{probability spaces on } \Omega\}$.
- For A : **Claim** define:

$$\begin{aligned} \mathcal{F}_A &:= \{(\Omega, \mathcal{S}, \mu) \mid A \in \mathcal{S}\} \\ P_A &:= \{(\Omega, \mathcal{S}, \mu) \mid \mu(A) = 1\} \\ P_{\textbf{Bel}(A)} &:= \{(\Omega, \mathcal{S}, \mu) \mid \mu(A) \geq t\}. \end{aligned}$$

* Can allow μ to be a capacity of order 2. Here I assume μ is a probability measure for simplicity.

Lockeans follow IPR

Theorem

Lockean semantics are sound for IPR.

Translation of syntax:

IPR	Probability	Interpretation
Δ **ctx**	$\Delta \subseteq \mathcal{P}(\Omega)$	The context for beliefs is determined by credences about claims.
$\Delta \vdash A$: **Claim**	$\Delta \subseteq \mathcal{F}_A$	A claim has meaning in context if it is assigned a credence.
$\Delta \vdash a : A$	$\Delta \subseteq P_A$	Assert that A is true if maximal credence in A.
$\Delta \vdash a$: **Bel**(A)	$\Delta \subseteq P_{\textbf{Bel}(A)}$	Assert belief in A if credence $\geq t$.

Sketch of Proof (for $A \times B$):

$$\frac{\Delta \vdash A : \textbf{Type} \qquad \Delta \vdash B : \textbf{Type}}{\Delta, a : A, b : B \vdash \texttt{pair}(a,b) : A \times B} \iff \frac{\Delta \subseteq \mathcal{F}_A \qquad \Delta \subseteq \mathcal{F}_B}{\Delta \cap P_A \cap P_B \subseteq P_{A \times B}}$$

Any $\mathcal{S}$ for which $A, B \in \mathcal{S}$ then $A \cap B \in \mathcal{S}$. If $P(A) = P(B) = 1$, then

$$P(A \cap B) = 1 - P(A^c \cup B^c) \geq 1 - (P(A^c) + P(B^c)) = 1 - ((1 - P(A)) + (1 - P(B))) = 1.$$

Conclusion

- Lockean semantics are sound for fragment of MLTT given by $\times$, $+$, **0**, **Bel**.

  H. Crane. (2019). Imprecise probabilities as a semantics for intuitive probabilistic reasoning. Researchers.One*, https://www.researchers.one/article/2018-08-8.*

- Interpretation of **Bel**-type in terms of typicality:

  Justified Belief in A $\iff$ A is typical.

  H. Crane and I. Wilhelm. (2019). The Logic of Typicality. In Valia Allori (ed.), Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism and Laws of Nature*, World Scientific.*

- Further technical details:

  H. Crane. (2018). Logic of Probability and Conjecture. Researchers.One*, https://www.researchers.one/article/2018-08-5*