Logics of Imprecise Comparative Probability

Wesley H. Holliday
University of California, Berkeley

Joint work with Yifeng Ding (Berkeley) and Thomas Icard (Stanford)

Logic Online Seminar
March 15, 2021

Outline

This talk is based on the paper "Logics of Imprecise Comparative Probability," forthcoming in *International Journal of Approximate Reasoning*.

- Comparative probability
- Imprecise probability
- Representation theorems
- Logics of imprecise comparative probability
  - Adding epistemic possibility
  - Dynamic logic of updating imprecise probability
  - Dynamic logic of becoming aware of a proposition
- Open problems

Comparative probability

One school of thought in the foundations of probability takes at its starting point a binary relation $\succsim$ between propositions:

$$A \succsim B \text{ means that } A \text{ is at least as likely as } B.$$

What axioms should $\succsim$ obey? Example:

$$\text{if } (A \cup B) \cap C = \varnothing, \text{ then } A \succsim B \text{ iff } (A \cup C) \succsim (B \cup C).$$

When is $\succsim$ representable by numerical measures of certain kinds, including Kolmogorov's probability measures?

Imprecise probability

Another school of thought critiques the assumption of the standard probability calculus that every proposition is assigned a precise probability value.

This school allow us to represent a state of uncertainty not just with a single probability measure but also with a *set* of probability measures.

Given a set $\mathcal{P}$ of probability measures and proposition A, an agent's attitude toward A is given by a set of possible probabilities

$$\{\mu(A) \mid \mu \in \mathcal{P}\}$$

or the associated interval

$$[\inf\{\mu(A) \mid \mu \in \mathcal{P}\}, \sup\{\mu(A) \mid \mu \in \mathcal{P}\}].$$

Imprecise comparative probability

Comparative Probability and Imprecise Probability can be related as follows.

Given a set $\mathcal{P}$ of probability measures, we define a binary relation by:

$$A \succsim B \text{ iff for all } \mu \in \mathcal{P}, \mu(A) \geq \mu(B).$$

We say that $\succsim$ is *represented by $\mathcal{P}$ as the weak relation*.

We can then ask what axioms on $\succsim$ are necessary and sufficient for the existence of a set $\mathcal{P}$ of probability measures that represents $\succsim$ as the weak relation.

Example

Consider the famous Ellsberg Paradox setup: in an urn of 100 marbles, 30 are known to be red, and the other 70 are blue or yellow in some unknown proportion.

People typically *prefer* a bet that yields \$1 if they draw a red marble over a bet that yields \$1 if they draw a blue one. This suggest that they do not *judge* that drawing a blue marble is more likely than drawing a red marble.

Yet with a single probability representation, the natural choice would be a measure with $\mu(R) = .3$ and $\mu(B) = \mu(Y) = .35$, so $B \succ R$.

Imprecise probability offers a different representation: the set $\mathcal{P} = \{\mu_n \mid n \in \{0, \ldots, 70\}\}$ where $\mu_n(R) = .3$, $\mu_n(B) = \frac{n}{100}$, $\mu_n(Y) = \frac{70-n}{100}$.

According to the imprecise representation, we have $R \not\succsim B$ and $B \not\succsim R$.

The decision rule of picking the bet whose lowest expected utility according to the measures in $\mathcal{P}$ is highest results in preferring the bet on R to the bet on B.

Representation

Theorem (Kraft et al. 1959)

Let W be a nonempty finite set and $\succsim$ a binary relation on $\wp(W)$. Then $\succsim$ is represented by some probability measure on $\wp(W)$ if and only if:

1. $\varnothing \not\succsim W$, *and* $\{w\} \succsim \varnothing$ *for all* $w \in W$*;*
2. *for all* $A, B \in \wp(W)$, $A \succsim B$ *or* $B \succsim A$*;*
3. $\succsim$ *satisfies the* **finite cancellation** *condition (FC): letting* $\mathbf{1}_X$ *be the characteristic function of* X*, for any two finite sequences* $\langle A_i \rangle_{i=1}^n, \langle B_i \rangle_{i=1}^n$ *of events in* $\wp(W)$ *such that*

$$\sum_{i=1}^{n} \mathbf{1}_{A_i} = \sum_{i=1}^{n} \mathbf{1}_{B_i}$$

(*additions are done in the vector space* $\mathbb{R}^W$),

$$\textit{if for all } i < n,\ A_i \succsim B_i, \textit{ then } B_n \succsim A_n.$$

Theorem (Ríos Insua 1992)

Let W be a nonempty finite set and $\succsim$ a binary relation on $\wp(W)$. Then $\succsim$ is represented as the weak relation by some set $\mathcal{P}$ of probability measures on $\wp(W)$ if and only if:

1. $\varnothing \not\succsim W$, *and* $\{w\} \succsim \varnothing$ *for all* $w \in W$;
2. $\succsim$ *satisfies the* **generalized finite cancellation condition** (GFC): *for any two finite sequences* $\langle A_i \rangle_{i=1}^n, \langle B_i \rangle_{i=1}^n$ *of events in* $\wp(W)$ *and* $k \in \mathbb{N} \setminus \{0\}$ *such that*
$$\sum_{i=1}^{n-1} \mathbf{1}_{A_i} + k\mathbf{1}_{A_n} = \sum_{i=1}^{n-1} \mathbf{1}_{B_i} + k\mathbf{1}_{B_n},$$
if for all $i < n$, $A_i \succsim B_i$, *then* $B_n \succsim A_n$.

Theorem (Harrison-Trainor, H., and Icard 2016)

There are relations $\succsim$ *satisfying the axioms in part 1 and* (FC) *but not* (GFC).

Adding $\succ$

Given a set $\mathcal{P}$ of measures, there are two ways to define a strict relation:

$$A \succnsim B \text{ iff } A \succsim B \text{ and } B \not\succsim A;$$
$$A \succ B \text{ iff for all } \mu \in \mathcal{P}, \mu(A) > \mu(B).$$

In the second case, we say that $\succ$ is *represented by* $\mathcal{P}$ *as the strict relation*.

E.g., suppose $\mathcal{P} = \{\mu_1, \mu_2\}$ with $\mu_1(A) = \mu_1(B)$ and $\mu_2(A) > \mu_2(B)$.

Then $A \succnsim B$ but not $A \succ B$.

Proposition

There are sets $\mathcal{P}, \mathcal{P}'$ whose weak relations $\succsim$ are the same but whose strict relations $\succ_{\mathcal{P}}$ and $\succ_{\mathcal{P}'}$ are different; and there are sets $\mathcal{P}, \mathcal{P}'$ whose strict relations $\succ$ are the same but whose weak relations $\succsim_{\mathcal{P}}$ and $\succsim_{\mathcal{P}'}$ are different.

A pair $\langle \succsim, \succ \rangle$ is represented by a set $\mathcal{P}$ if $\succsim$ is represented as the weak relation by $\mathcal{P}$ and $\succ$ is represented as the strict relation by $\mathcal{P}$.

Theorem (Ding, H., and Icard 2021)

Let W be a nonempty finite set and $\succsim, \succ$ two binary relations on $\wp(W)$. Then $\langle \succsim, \succ \rangle$ is represented by a set $\mathcal{P}$ of probability measures on $\wp(W)$ if and only if:

- *$\succ$ is irreflexive and $\succ \subseteq \succsim$;*
- *$W \succ \varnothing$, and $\{w\} \succsim \varnothing$ for all $w \in W$;*
- *$\succsim$ satisfies (GFC) and $\succ$ satisfies the* **strict generalized finite cancellation condition** *$(SGFC)$: for any two finite sequences $\langle A_i \rangle_{i=1}^{n}, \langle B_i \rangle_{i=1}^{n}$ of events in $\wp(W)$ and $k \in \mathbb{N} \setminus \{0\}$ such that*

$$\sum_{i=1}^{n-1} \mathbf{1}_{A_i} + k\mathbf{1}_{A_n} = \sum_{i=1}^{n-1} \mathbf{1}_{B_i} + k\mathbf{1}_{B_n},$$

if for all $i < n$, $A_i \succsim B_i$, and there is $i < n$ with $A_i \succ B_i$, then $B_n \succ A_n$.

The Logic IP($\succsim, \succ$)

The language $\mathcal{L}(\succsim, \succ)$, generated from a nonempty set Prop of propositional variables, is defined by the following grammar, where $p \in \text{Prop}$:

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid (\varphi \succsim \varphi) \mid (\varphi \succ \varphi).$$

The semantics is given in terms of imprecise probability as expected:

- $\mathcal{M}, \mathcal{P}, w \vDash \varphi \succsim \psi$ iff for all $\mu \in \mathcal{P}$, $\mu([\![\varphi]\!]^{\mathcal{M},\mathcal{P}}) \geq \mu([\![\psi]\!]^{\mathcal{M},\mathcal{P}})$.
- $\mathcal{M}, \mathcal{P}, w \vDash \varphi \succ \psi$ iff for all $\mu \in \mathcal{P}$, $\mu([\![\varphi]\!]^{\mathcal{M},\mathcal{P}}) > \mu([\![\psi]\!]^{\mathcal{M},\mathcal{P}})$.

We then define a logic IP($\succsim, \succ$) with axioms corresponding to the conditions of the representation theorem for pairs $\langle \succsim, \succ \rangle$.

Theorem (Ding, H., and Icard 2021)

IP($\succsim, \succ$) *is sound and complete with respect to imprecise probability semantics.*

Adding epistemic possibility

Next we add a "possibility" modal $\diamond$ to the language. In addition to making claims about the possibility of factual states of affairs, e.g.,

"It is possible that it is raining,"

we want to make claims about the possibility of likelihood relations, e.g.,

"It is possible that hail is more likely than lightning tonight."

The semantics for $\diamond$ is:

- $\mathcal{M}, \mathcal{P}, w \vDash \diamond\varphi$ iff there is a $\mu \in \mathcal{P}$ such that $\mu(\llbracket\varphi\rrbracket^{\mathcal{M},\{\mu\}}) \neq 0$.

Theorem (Ding, H., and Icard 2021)

The complexity of the satisfiability problem for $\mathcal{L}(\succsim, \succ, \diamond)$ is NP-complete.

We also give a sound and complete logic $\mathsf{IP}(\succsim, \succ, \diamond)$.

Dynamic logic of updating imprecise probability

Given an initial set $\mathcal{P}$ of probability measures, after learning some proposition $U \subseteq W$ with certainty, we update the set $\mathcal{P}$ to the set

$$\mathcal{P}_U = \{\mu(\cdot \mid U) : \mu \in \mathcal{P}, \mu(U) > 0\},$$

where $\mu(\cdot \mid U)$ is defined as usual: for any $V \subseteq W$, $\mu(V \mid U) = \frac{\mu(V \cap U)}{\mu(U)}$.

Since we have a formal language with comparative probability operators, we can model updating on sentences containing not only factual formulas but also comparative probability formulas, as in "it is raining, and it is more likely that there will be hail than it is that there will be lightning" $(r \wedge (h \succ \ell))$

Dynamic logic of updating imprecise probability

As in dynamic epistemic logic, we add dynamic operators whose semantics is given in terms of the update operation:

- $\mathcal{M}, \mathcal{P}, w \vDash \langle \varphi \rangle \psi$ iff there is a $\mu \in \mathcal{P}$ such that $\mu(\llbracket \varphi \rrbracket^{\mathcal{M}, \{\mu\}}) \neq 0$ and $\mathcal{M}, \mathcal{P}_\varphi, w \vDash \psi$,

where

$$\mathcal{P}_\varphi = \{\nu(\cdot \mid \llbracket \varphi \rrbracket^{\mathcal{M}, \{\nu\}}) : \nu \in \mathcal{P} \text{ and } \nu(\llbracket \varphi \rrbracket^{\mathcal{M}, \{\nu\}}) \neq 0\}.$$

If $\varphi := r \wedge (h \succ \ell)$, then $\langle \varphi \rangle$ has the effect of discarding all measures that do not satisfy $h \succ \ell$ and then conditioning each of the remaining measures on r.

Dynamic logic of updating imprecise probability

The logic $\mathsf{IP}(\succsim, \succ, \Diamond, \langle\,\rangle)$ is the smallest set of $\mathcal{L}(\succsim, \succ, \Diamond, \langle\,\rangle)$ formulas that is (i) closed under modus ponens and the rule of replacement of equivalents, and (ii) contains all theorems of $\mathsf{IP}(\succsim, \succ, \Diamond)$ as well as all instances of the following axiom schemas where $p \in \mathsf{Prop}$ and α and β are propositional:

(**R0**) $\langle\varphi\rangle p \leftrightarrow (\Diamond\varphi \wedge p)$;

(**R1**) $\langle\varphi\rangle\Diamond\psi \leftrightarrow \Diamond\langle\varphi\rangle\psi$;

(**R2**) $\langle\varphi\rangle\neg\psi \leftrightarrow (\Diamond\varphi \wedge \neg\langle\varphi\rangle\psi)$;

(**R3**) $\langle\varphi\rangle(\psi \wedge \chi) \leftrightarrow (\langle\varphi\rangle\psi \wedge \langle\varphi\rangle\chi)$;

(**R4**) $\langle\varphi\rangle(\alpha \succsim \beta) \leftrightarrow (\Diamond\varphi \wedge \Box((\varphi \wedge \alpha) \succsim (\varphi \wedge \beta)))$;

(**R5**) $\langle\varphi\rangle(\alpha \succ \beta) \leftrightarrow (\Diamond\varphi \wedge \Box((\varphi \succ \bot) \rightarrow ((\varphi \wedge \alpha) \succ (\varphi \wedge \beta))))$.

Theorem (Ding, H., and Icard 2021)

$\mathsf{IP}(\succsim, \succ, \Diamond, \langle\,\rangle)$ *is sound and complete.*

Dynamic logic of updating imprecise probability

The logic $\mathsf{IP}(\succsim, \succ, \Diamond, \langle\,\rangle)$ is the smallest set of $\mathcal{L}(\succsim, \succ, \Diamond, \langle\,\rangle)$ formulas that is (i) closed under modus ponens and the rule of replacement of equivalents, and (ii) contains all theorems of $\mathsf{IP}(\succsim, \succ, \Diamond)$ as well as all instances of the following axiom schemas where $p \in \mathsf{Prop}$ and α and β are propositional:

(R0) $\langle\varphi\rangle p \leftrightarrow (\Diamond\varphi \wedge p)$;

(R1) $\langle\varphi\rangle\Diamond\psi \leftrightarrow \Diamond\langle\varphi\rangle\psi$;

(R2) $\langle\varphi\rangle\neg\psi \leftrightarrow (\Diamond\varphi \wedge \neg\langle\varphi\rangle\psi)$;

(R3) $\langle\varphi\rangle(\psi \wedge \chi) \leftrightarrow (\langle\varphi\rangle\psi \wedge \langle\varphi\rangle\chi)$;

(R4) $\langle\varphi\rangle(\alpha \succsim \beta) \leftrightarrow (\Diamond\varphi \wedge \Box((\varphi \wedge \alpha) \succsim (\varphi \wedge \beta)))$;

(R5) $\langle\varphi\rangle(\alpha \succ \beta) \leftrightarrow (\Diamond\varphi \wedge \Box((\varphi \succ \bot) \rightarrow ((\varphi \wedge \alpha) \succ (\varphi \wedge \beta))))$.

There is exponential blowup in the size of formulas when applying these axioms to translate $\mathcal{L}(\succsim, \succ, \Diamond, \langle\,\rangle)$ to $\mathcal{L}(\succsim, \succ, \Diamond)$. Is there a satisfiability-preserving translation with only polynomial blowup (as Lutz 2006 showed for PAL)?

Becoming aware of a proposition

Example

A patient learns from her doctor of the existence of a gland in the human body and of a disease previously unknown to her.[1] The doctor informs her that if her gland is swollen, then it is more likely than not that she has the disease. Subsequently the patient's gland is examined, and she learns that it is swollen. As a result, she comes to think it is more likely than not that she has the disease.

How should we model the patient's becoming aware of the gland and the disease?

[1] This example is inspired by van Benthem's (2011, p. 164) example of the hypochondriac.

Dynamic logic of becoming aware of a proposition

The language $\mathcal{L}(\succsim, \succ, \Diamond, \langle\ \rangle, I)$ is defined by the following grammar:

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid (\varphi \succsim \varphi) \mid (\varphi \succ \varphi) \mid \Diamond\varphi \mid \langle\varphi\rangle\varphi \mid I_p^+\varphi \mid I_p^-\varphi.$$

- read $I_p^+\varphi$ as "letting p be a true proposition that is newly introduced to the agent, φ";
- read $I_p^-\varphi$ as "letting p be a false proposition that is newly introduced to the agent, φ";
- take $I_p\varphi$ as an abbreviation of $(I_p^+\varphi \wedge I_p^-\varphi)$.

When introducing a new p, we “split” each $w \in W$ into a world $\langle w, 0\rangle$ where p is false and a world $\langle w, 1\rangle$ where p is true. So we move to $W \times \{0, 1\}$. Let $\mathcal{F}$ be a field of sets on W and V a valuation such that $V(q) \in \mathcal{F}$ for all $q \in \mathsf{Prop}$.

- V^{+p} is defined by $V^{+p}(q) = \begin{cases} V(q) \times \{0,1\} & \text{if } q \neq p \\ W \times \{1\} & \text{if } q = p\,. \end{cases}$
- Given $\mathcal{M} = \langle W, V\rangle$, let $\mathcal{M}^{+p} = \langle W \times \{0,1\}, V^{+p}\rangle$.
- $\mathcal{F} \times 2 = \{X \times \{0,1\} \mid X \in \mathcal{F}\}$, which is a field of sets on $W \times \{0,1\}$.
- $\mathsf{Split}(\mathcal{F})$ is the smallest field on $W \times \{0,1\}$ extending $\mathcal{F} \cup \{W \times \{0\}\}$.
- For any finitely additive measure μ on $\mathcal{F}$, define $\mu \times 2$, a finitely additive measure on $\mathcal{F} \times 2$, by $\mu \times 2(X \times \{0,1\}) = \mu(X)$ for all $X \in \mathcal{F}$.
- $\mathcal{P} \times 2 = \{\mu \times 2 \mid \mu \in \mathcal{P}\}$.
- $\mathsf{Split}(\mathcal{P})$ is the set of all finitely additive measures μ on $\mathsf{Split}(\mathcal{F})$ such that $\mu|_{\mathcal{F}\times 2} \in \mathcal{P} \times 2$.

Dynamic logic of becoming aware of a proposition

The semantics of I_p^+ and I_p^- are given by

$$\mathcal{M}, \mathcal{P}, w \models I_p^- \varphi \text{ iff } \mathcal{M}^{+p}, \mathsf{Split}(\mathcal{P}), \langle w, 0 \rangle \models \varphi\,,$$
$$\mathcal{M}, \mathcal{P}, w \models I_p^+ \varphi \text{ iff } \mathcal{M}^{+p}, \mathsf{Split}(\mathcal{P}), \langle w, 1 \rangle \models \varphi\,.$$

The following sentence is valid and represents the medical example if we take the proposition introduced by I_p to be that the agent has the disease and the proposition introduced by I_q to be that the gland is swollen:

$$I_p \langle \neg p \succ p \rangle I_q \langle (q \wedge p) \succ (q \wedge \neg p) \rangle \langle q \rangle (p \succ \neg p).$$

We interpret the first update by $\neg p \succ p$ as the result of the agent observing that she is not feeling uncomfortable and hence believing that her not having the disease is more likely than her having it.

Expressivity

The language with I_p^+ and I_p^- allows us to express comparisons of linear combinations of probabilities with rational coefficients.

Basic idea: introduce new propositions standing for independent coin flips coming up heads. E.g., the following says that the probability of p is greater than $3/4$:

$I_q I_r \langle ((q \wedge r) \approx (q \wedge \neg r)) \wedge ((q \wedge r) \approx (\neg q \wedge r)) \wedge ((q \wedge r) \approx (\neg q \wedge \neg r)) \rangle (p \succ \neg(q \wedge r))$.

Proposition

For any sequences $\langle \varphi_i \rangle_{i=1\dots n}$ *and* $\langle \psi_i \rangle_{i=1\dots m}$ *of formulas in* $\mathcal{L}(\succsim, \succ, \Diamond, \langle\ \rangle, I)$ *and any sequences* $\langle a_i \rangle_{i=1\dots n}$ *and* $\langle b_i \rangle_{i=1\dots m}$ *of natural numbers, there is a formula* $\chi \in \mathcal{L}(\succsim, \succ, \Diamond, \langle\ \rangle, I)$ *such that for any IP model* $\mathcal{M}, \mathcal{P}, w$,

$$\mathcal{M}, \mathcal{P}, w \models \chi \textit{ iff } \forall \mu \in \mathcal{P}, \sum_{i=1}^{n} a_i \mu(\llbracket \varphi_i \rrbracket^{\mathcal{M},\mathcal{P}}) \geq \sum_{i=1}^{m} b_i \mu(\llbracket \psi_i \rrbracket^{\mathcal{M},\mathcal{P}}).$$

Open problems

Problem

Find an axiomatization of the set of valid formulas in $\mathcal{L}(\succsim, \succ, \Diamond, \langle\,\rangle, I)$.

Problem

Determine the complexity of the satisfiability problem for $\mathcal{L}(\succsim, \succ, \Diamond, \langle\,\rangle, I)$.

There is also the open problem for our less expressive dynamic language:

Problem

Determine the complexity of the satisfiability problem for $\mathcal{L}(\succsim, \succ, \Diamond, \langle\,\rangle)$.

Conclusion

- Comparative probability remains a rich area for logical study.
- When teaching Introduction to Logic, we could introduce $\succsim$ right after introducing $\neg$, $\wedge$, $\vee$, etc., for propositional logic.
- We could take Boolean algebras with $\succsim$ as objects for universal-algebraic study, develop representation and duality theory, etc.
- We hope to see fruitful new contacts between logic and probability.

Thank you!