HOL Theorem Proving and Formal Probability (2)

Chun TIAN
chun.tian@anu.edu.au

20/03/2024

Outline of Contents

- Preliminary (set theory, topology, reals and extended reals, etc.)
- Measure Theory
- Borel Measure Space
- Lebesgue Integration Theory
- Probability Theory
- Simple Stochastic Processes
  - Random sequences (IID and stationary)
  - Martingale
  - (Markov chain)

History of HOL4-Probability

- *Joe Hurd*, **Formal verification of probabilistic algorithms,** University of Cambridge, UCAM-CL-TR-566, 2003.
- *Osman Hasan*, **Formal Probabilistic Analysis using Theorem Proving,** Concordia University (Hardware Verification Group), 2008.
- *Aaron R. Coble*, **Anonymity, information, and machine-assisted proof,** University of Cambridge, UCAM-CL-TR-785, 2010.
- *Tarek Mhamdi*, **Information-Theoretic Analysis Using Theorem Proving,** Concordia University (Hardware Verification Group), 2012.
- *Chun Tian*, **Carathéodory extension theorem for semirings**, **Construction of the one-dimensional Borel measure**, **Product measure**, **The Law of Large Numbers**, etc. 2018-2022 (no publications)

Reference Books

- A. N. Kolmogorov, Foundations of the Theory of Probability (Grundbegriffe der Wahrscheinlichkeitsrechnung). Chelsea Publishing Company, 1950. (orig. 1933)
- R. L. Schilling, Measures, Integrals and Martingales, 2nd ed., Cambridge University Press, 2017.
- K. L. Chung, A Course in Probability Theory, 3rd ed., Academic Press, 2001.
- J. S. Rosenthal, A First Look at Rigorous Probability Theory, 2nd ed., World Scientific Publishing Company, 2006.
- P. Billingsley, Probability and Measure, 3rd ed., John Wiley & Sons, 1995.
- A. N. Shiryaev, Probability-1, 3rd ed., Springer-Verlag, 2016. (orig. 2007)
- A. N. Shiryaev, Probability-2, 3rd ed., Springer-Verlag, 2019.

Family of Sets: Algebra (aka Field)

$(X, \mathcal{A})$ is an *algebra* (or field) if:

- $A \in \mathcal{A} \Rightarrow A \subseteq X$
- $\emptyset \in \mathcal{A}$
- $A \in \mathcal{A} \Rightarrow X \setminus A \in \mathcal{A}$ (or $\overline{A} \in \mathcal{A}$)
- $A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$

$$
\begin{array}{l}
\vdash \texttt{subset\_class}\ sp\ sts \iff \forall x.\ x \in sts \Rightarrow x \subseteq sp \\
\vdash \texttt{algebra}\ a \iff \\
\quad \texttt{subset\_class}\ (\texttt{space}\ a)\ (\texttt{subsets}\ a) \wedge \\
\quad \emptyset \in \texttt{subsets}\ a \wedge \\
\quad (\forall s.\ s \in \texttt{subsets}\ a \Rightarrow \texttt{space}\ a\ \texttt{DIFF}\ s \in \texttt{subsets}\ a) \wedge \\
\quad \forall s\ t. \\
\qquad s \in \texttt{subsets}\ a \wedge t \in \texttt{subsets}\ a \Rightarrow \\
\qquad s \cup t \in \texttt{subsets}\ a \qquad\qquad [\texttt{algebra\_def}]
\end{array}
$$

- $X \in \mathcal{A}$ (as $X = \overline{\emptyset}$ or $X \setminus \emptyset$)
- $A, B \in \mathcal{A} \Rightarrow A \cap B \in \mathcal{A}$ (as $\overline{A \cap B} = \overline{A} \cup \overline{B} \in \mathcal{A}$)
- $A, B \in \mathcal{A} \Rightarrow A \setminus B \in \mathcal{A}$ (as $A \setminus B = A \cap \overline{B} \in \mathcal{A}$)

System of Sets: σ-Algebra

$(X, \mathcal{A})$ is an *σ-algebra* if $(X, \mathcal{A})$ is algebra and, additionally:

$$\forall A_1, A_2, \ldots \in \mathcal{A} \Rightarrow \bigcup_{i \in \mathbb{N}} A_i \in \mathcal{A}$$

$$
\begin{array}{ll}
\vdash \texttt{sigma\_algebra}\ a \iff & \\
\quad \texttt{algebra}\ a\ \wedge & \\
\quad \forall c.\ \texttt{countable}\ c\ \wedge\ c \subseteq \texttt{subsets}\ a \Rightarrow & \\
\qquad \bigcup c \in \texttt{subsets}\ a & \texttt{[sigma\_algebra\_def]}
\end{array}
$$

$$
\begin{array}{ll}
\vdash \texttt{sigma\_algebra}\ (sp, sts) \iff & \\
\quad \texttt{algebra}\ (sp, sts)\ \wedge & \\
\quad \forall A.\ \texttt{IMAGE}\ A\ \mathcal{U}(\texttt{:num}) \subseteq sts \Rightarrow & \\
\qquad \bigcup \{A\ i \mid i \in \mathcal{U}(\texttt{:num})\} \in sts & \texttt{[sigma\_algebra\_alt]}
\end{array}
$$

Alternatively, $\forall A_1, A_2, \ldots \in \mathcal{A} \Rightarrow \bigcap_{i \in \mathbb{N}} A_i \in \mathcal{A}$

$$
\begin{array}{l}
\vdash \texttt{sigma\_algebra}\ (sp, sts)\ \wedge\ \texttt{countable}\ X\ \wedge\ X \neq \emptyset\ \wedge \\
\quad \texttt{IMAGE}\ A\ X \subseteq sts \Rightarrow \\
\quad \bigcap \{A\ x \mid x \in X\} \in sts
\end{array}
$$

Trivial (σ-)algebras

Some trivial algebras (and also σ-algebras):

- Smallest (σ-)algebra generated from any space X: $(X, \{\emptyset, X\})$
- Single-set (σ-)algebra: $S \subseteq X \Rightarrow (X, \{\emptyset, S, X \setminus S, X\})$
- Biggest (σ-)algebra generated from any space X: $(X, \mathcal{P}(X))$

In particular, finite algebra (finite space or finite subsets) is always σ-algebra:

```
[algebra_finite_space_imp_sigma_algebra]
 ⊢ algebra a ∧ FINITE (space a) ⇒ sigma_algebra a

[algebra_finite_subsets_imp_sigma_algebra]
 ⊢ algebra a ∧ FINITE (subsets a) ⇒
   sigma_algebra a
```

Extended Real Numbers

`extreal = NegInf | PosInf | Normal real`

[extreal_add_def]
$\vdash$ Normal x + Normal y = Normal $(x + y)$ $\wedge$
Normal v_0 + $-\infty$ = $-\infty$ $\wedge$ Normal v_0 + $+\infty$ = $+\infty$ $\wedge$
$-\infty$ + Normal v_1 = $-\infty$ $\wedge$ $+\infty$ + Normal v_1 = $+\infty$ $\wedge$
$-\infty + -\infty = -\infty \wedge +\infty + +\infty = +\infty$
[extreal_ainv_def]
$\vdash --\infty = +\infty \wedge -+\infty = -\infty \wedge \forall x.$ $-$Normal x = Normal $(-x)$
[extreal_sub]
$\vdash x - y = x + -y$
[extreal_div]
$\vdash$ ($\forall r.$ Normal r / $+\infty$ = Normal 0) $\wedge$
($\forall r.$ Normal r / $-\infty$ = Normal 0) $\wedge$
$\forall x\ r.\ r \neq 0 \Rightarrow x$ / Normal r = $x \times$ (Normal r)$^{-1}$
[extreal_inv_def]
$\vdash -\infty^{-1}$ = Normal 0 $\wedge$ $+\infty^{-1}$ = Normal 0 $\wedge$
$\forall r.\ r \neq 0 \Rightarrow$ (Normal r)$^{-1}$ = Normal r^{-1}

NOTE: $\infty + -\infty$, $\infty - \infty$, $\frac{\infty}{\infty}$ and division-by-zero are unspecified.

Extended Real Numbers (2)

However, $0 \cdot \pm\infty = 0$ is defined to be 0 (consider a limiting process):

$\vdash -\infty \times -\infty = +\infty \wedge -\infty \times +\infty = -\infty \wedge +\infty \times -\infty = -\infty \wedge +\infty \times +\infty$
$\quad$ Normal $x \times -\infty$ =
$\quad$ (**if** $x = 0$ **then** Normal 0
$\quad$ **else if** $0 < x$ **then** $-\infty$
$\quad$ **else** $+\infty$) $\wedge$
$\quad$ $-\infty \times$ Normal y =
$\quad$ (**if** $y = 0$ **then** Normal 0
$\quad$ **else if** $0 < y$ **then** $-\infty$
$\quad$ **else** $+\infty$) $\wedge$
$\quad$ Normal $x \times +\infty$ =
$\quad$ (**if** $x = 0$ **then** Normal 0
$\quad$ **else if** $0 < x$ **then** $+\infty$
$\quad$ **else** $-\infty$) $\wedge$
$\quad$ $+\infty \times$ Normal y =
$\quad$ (**if** $y = 0$ **then** Normal 0
$\quad$ **else if** $0 < y$ **then** $+\infty$
$\quad$ **else** $-\infty$) $\wedge$ Normal $x \times$ Normal y = Normal $(x \times y)$ [extreal_mul_def]

$\vdash 0 \times x = 0$ [mul_lzero]
$\vdash x \times 0 = 0$ [mul_rzero]

Measure Space

A *measure space* is a tuple $(X, \mathcal{A}, \mu)$ (space, measurable sets, measure):

- $(X, \mathcal{A})$ (*measurable space*) is a σ-algebra;
- $\mu : \mathcal{A} \to [0, +\infty]$ is a *measure* such that (cf. premeasure)
  - $\mu(\emptyset) = 0$
  - $\forall s \in \mathcal{A}.\, 0 \leqslant \mu(s)$
  - $A_i \cap A_j = \emptyset \Rightarrow \mu(A_1 \cup A_2 \cdots) = \sum_{i \in \mathbb{N}} \mu(A_i)$

$\vdash$ measure_space m $\iff$
  sigma_algebra (measurable_space m) $\wedge$
  positive m $\wedge$ countably_additive m [measure_space_def]

$\vdash$ positive m $\iff$
  measure m $\emptyset$ = 0 $\wedge$
  $\forall s.$ $s \in$ measurable_sets m $\Rightarrow$ $0 \leq$ measure m s [positive_def]

$\vdash$ countably_additive m $\iff$
  $\forall f.$ $f \in$ ($\mathcal{U}$(:num) $\to$ measurable_sets m) $\wedge$
    ($\forall i\ j.$ $i \neq j$ $\Rightarrow$ DISJOINT (f i) (f j)) $\wedge$
    $\bigcup$ (IMAGE f $\mathcal{U}$(:num)) $\in$ measurable_sets m $\Rightarrow$
    measure m ($\bigcup$ (IMAGE f $\mathcal{U}$(:num))) =
    suminf (measure m $\circ$ f) [countably_additive_def]

More Properties of (pre)measure

$$
\begin{array}{l}
\vdash \texttt{increasing}\ m \iff \\
\quad \forall s\ t.\\
\qquad s \in \texttt{measurable\_sets}\ m \land \\
\qquad t \in \texttt{measurable\_sets}\ m \land s \subseteq t \Rightarrow \\
\qquad \texttt{measure}\ m\ s \le \texttt{measure}\ m\ t \\
\vdash \texttt{additive}\ m \iff \\
\quad \forall s\ t.\\
\qquad s \in \texttt{measurable\_sets}\ m \land \\
\qquad t \in \texttt{measurable\_sets}\ m \land \texttt{DISJOINT}\ s\ t \land \\
\qquad s \cup t \in \texttt{measurable\_sets}\ m \Rightarrow \\
\qquad \texttt{measure}\ m\ (s \cup t) = \texttt{measure}\ m\ s + \texttt{measure}\ m\ t \\
\vdash \texttt{subadditive}\ m \iff \\
\quad \forall s\ t.\\
\qquad s \in \texttt{measurable\_sets}\ m \land \\
\qquad t \in \texttt{measurable\_sets}\ m \land \\
\qquad s \cup t \in \texttt{measurable\_sets}\ m \Rightarrow \\
\qquad \texttt{measure}\ m\ (s \cup t) \le \texttt{measure}\ m\ s + \texttt{measure}\ m\ t \\
\vdash \texttt{countably\_subadditive}\ m \iff \\
\quad \forall f.\ f \in (\mathcal{U}(\texttt{:num}) \to \texttt{measurable\_sets}\ m) \land \\
\qquad\quad \bigcup\ (\texttt{IMAGE}\ f\ \mathcal{U}(\texttt{:num})) \in \texttt{measurable\_sets}\ m \Rightarrow \\
\qquad\quad \texttt{measure}\ m\ (\bigcup\ (\texttt{IMAGE}\ f\ \mathcal{U}(\texttt{:num}))) \le \\
\qquad\quad \texttt{suminf}\ (\texttt{measure}\ m \circ f)
\end{array}
$$

Monotone Convergence of (pre)measure

$$f_n \subseteq f_{n+1} \in \mathcal{A} \Rightarrow \sup\{\mu(f_i)\} = \mu(\bigcup_{i \in \mathbb{N}} f_i)$$

[MONOTONE_CONVERGENCE2]
$\vdash$ `measure_space` m $\wedge$
  f $\in$ ($\mathcal{U}$(`:num`) $\to$ `measurable_sets` m) $\wedge$
  ($\forall\, n$. f n $\subseteq$ f (`SUC` n)) $\Rightarrow$
  `sup (IMAGE (measure` $m \circ f$) $\mathcal{U}$(`:num`)) =
  `measure` m ($\bigcup$ (`IMAGE` f $\mathcal{U}$(`:num`)))

$$f_{n+1} \subseteq f_n \in \mathcal{A} \Rightarrow \inf\{\mu(f_i)\} = \mu(\bigcap_{i \in \mathbb{N}} f_i) \qquad (\mu(f_i) \neq \infty)$$

[MONOTONE_CONVERGENCE_BIGINTER2]
$\vdash$ `measure_space` m $\wedge$
  f $\in$ ($\mathcal{U}$(`:num`) $\to$ `measurable_sets` m) $\wedge$
  ($\forall\, n$. `measure` m (f n) $\neq$ $+\infty$) $\wedge$
  ($\forall\, n$. f (`SUC` n) $\subseteq$ f n) $\Rightarrow$
  `inf (IMAGE (measure` $m \circ f$) $\mathcal{U}$(`:num`)) =
  `measure` m ($\bigcap$ (`IMAGE` f $\mathcal{U}$(`:num`)))

Probability Space

A measure space $(X, \mathcal{A}, \mu)$ is called *probability space* if $\mu(X) = 1$.

$\vdash$ `prob_space` p $\iff$
`measure_space` p $\wedge$ `measure` p `(m_space` p`) = 1` [prob_space_def]

Probability space is usually denoted by $(\Omega, \mathcal{A}, P)$, where:

- Ω is called the *sample space*;
- $\mathcal{A}$ is called the *set of events* and $E \in \mathcal{A}$ a *event*;
- $P(E)$ is called the *probability* of event E.

$\vdash$ `prob_space` p $\iff$
`sigma_algebra (p_space` p`,events` p`)` $\wedge$
`positive` p $\wedge$ `countably_additive` p $\wedge$
`prob` p `(p_space` p`) = 1` [PROB_SPACE]

A sample probability space (for one-time coin tossing):
$(\{H, T\}, \{\emptyset, \{H\}, \{T\}, \{H, T\}\}, P)$, where $P\{H\} = P\{T\} = 0.5$.

Independent Events; Conditional Probability

Two events A and B are *independent* if $P(A \cap B) = P(A)\ P(B)$:

$\vdash$ `indep` p a b $\iff$
  $a \in$ `events` p $\wedge$ $b \in$ `events` p $\wedge$
  `prob` p ($a \cap b$) = `prob` p a $\times$ `prob` p b [indep_def]

$$P(E_1|E_2) := \frac{P(E_1 \cap E_2)}{P(E_2)} \quad \text{(conditional probability)}$$

$\vdash$ `cond_prob` p e_1 e_2 =
  `prob` p ($e_1 \cap e_2$) / `prob` p e_2 [cond_prob_def]

Bayes' formula:

$$P(B|A) = P(A|B) \cdot \frac{P(B)}{P(A)} \qquad (P(A) \neq 0, P(B) \neq 0)$$

$\vdash$ `prob_space` p $\wedge$ $A \in$ `events` p $\wedge$ $B \in$ `events` p $\wedge$
  `prob` p $A \neq 0$ $\wedge$ `prob` p $B \neq 0$ $\Rightarrow$
  `cond_prob` p B A =
  `cond_prob` p A B $\times$ `prob` p B / `prob` p A [BAYES_RULE]

Exercises

1. Prove $(X, \{\emptyset, X\})$ is (σ-)algebra.

2. Prove $(\{H, T\}, \{\emptyset, \{H\}, \{T\}, \{H, T\}\}$ is (σ-)algebra.

3. Prove $(\{H, T\}, \{\emptyset, \{H\}, \{T\}, \{H, T\}\}, P)$, where $P\{H\} = P\{T\} = 0.5$, is indeed a probability space.

Some useful theorems:

```
[sigma_algebraTheory.algebra_finite_space_imp_sigma_algebra]
 ⊢ algebra a ∧ FINITE (space a) ⇒ sigma_algebra a

[measureTheory.finite_additivity_sufficient_for_finite_spaces2]
 ⊢ sigma_algebra (measurable_space m) ∧
   FINITE (m_space m) ∧ positive m ∧ additive m ⇒
   measure_space m
```