GandALF: probabilistic models

Jan Otop

May 10, 2018

Plan for today

- Probability crash course.
- Markov chains.
- Markov Decision Processes.

Need for probability theory

- Probability $\mathrm{P}\colon 2^{\Omega} \to [0, 1]$.
- Probability over infinite words $\Omega = \Sigma^{\omega}$.
- Problem: only countable many words can have positive probability.

Need for probability theory

- Probability $\mathrm{P}\colon 2^\Omega \to [0,1]$.
- Probability over infinite words $\Omega = \Sigma^\omega$.
- Problem: only countable many words can have positive probability.

Problem: Vitali set in $[0,1]$.

Need for probability theory

- Probability $\mathrm{P}\colon 2^{\Omega} \to [0,1]$.
- Probability over infinite words $\Omega = \Sigma^{\omega}$.
- Problem: only countable many words can have positive probability.

Problem: Vitali set in $[0,1]$.

Definition

A triple $(\Omega, \mathcal{F}, \mathrm{P})$ is a probability space, if

- $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ such that $\mathcal{F}$ is a σ-field:
  - $\emptyset, \Omega \in \mathcal{F}$,
  - $\mathcal{F}$ is closed under complements, and
  - $\mathcal{F}$ is closed under countable unions
- $\mathrm{P}\colon \mathcal{F} \to [0,1]$ such that
  - $\mathrm{P}(\Omega) = 1$, and
  - P is countably additive, i.e., for disjoint $\mathrm{A}_1, \mathrm{A}_2 \ldots$
    $\mathrm{P}(\bigcup_{i \geq 1} \mathrm{A}_i) = \sum_{i=1}^{\infty} \mathrm{A}_i$.

Probabilistic space on Σ^ω

- Basic sets: $B = \{u \cdot \Sigma^\omega : u \in \Sigma^*\}$.
- The least σ-field $\mathcal{F}_B$ containing B — Borel sets.

Examples of Borel sets:

- $A = \{a^\omega\}$.
- $A_n = \{w \mid \text{at most n letters a in w}\}$.
- $A_{<\infty} = \{w \mid \text{finitely many letters a in w}\}$.
- $A_{even} = \{w \mid \text{w has letter a at every even position }\}$.

Probabilistic space on Σ^ω

- Basic sets: $B = \{u \cdot \Sigma^\omega : u \in \Sigma^*\}$.
- The least σ-field $\mathcal{F}_B$ containing B — Borel sets.

Examples of Borel sets:

- $A = \{a^\omega\}$.
- $A_n = \{w \mid \text{at most n letters a in w}\}$.
- $A_{<\infty} = \{w \mid \text{finitely many letters a in w}\}$.
- $A_{even} = \{w \mid \text{w has letter a at every even position }\}$.

All ω-regular sets are Borel.

Probabilistic space on Σ^ω

- Basic sets: $\mathrm{B} = \{\mathrm{u} \cdot \Sigma^\omega : \mathrm{u} \in \Sigma^*\}$.
- The least σ-field $\mathcal{F}_\mathrm{B}$ containing B — Borel sets.

Pre-measure — measure defined on B (weakly σ-additive).

Theorem (Carathéodory's extension theorem (specialized))

Any probability pre-measure μ_0 defined on ring(B), can be extended to a measure μ on $\mathcal{F}_\mathrm{B}$.

Example: Probability defined on basic sets, can be extended to $\mathcal{F}_\mathrm{B}$.

Examples

Consider $\Sigma = \{a, b\}$ and μ be such that $P(u\Sigma^\omega) = 2^{-|u|}$.

- $A = \{a^\omega\}$.
- $A_n = \{w \mid \text{at most n letters a in w}\}$.
- $A_{<\infty} = \{w \mid \text{finitely many letters a in w}\}$.
- $A_{even} = \{w \mid \text{w has letter a at every even position }\}$.

Examples

Consider $\Sigma = \{a, b\}$ and μ be such that $P(u\Sigma^\omega) = 2^{-|u|}$.

- $A = \{a^\omega\}$.
- $A_n = \{w \mid \text{at most n letters a in w}\}$.
- $A_{<\infty} = \{w \mid \text{finitely many letters a in w}\}$.
- $A_{even} = \{w \mid \text{w has letter a at every even position}\}$.

Kolmogorov 0-1 law — if A is prefix-independent, then $P(A) = 0$ or $P(A) = 1$.

Markov chains

Definition

A Markov chain is a tuple (S, s_0, E) such that S is a finite set of states, and $E\colon S \times S \to [0, 1]$, and for each s we have $\sum_{t \in S} E(s, t) = 1$.

- Labeled Markov chains.
- Generate probability space on S^ω.
- Matrix representation.

Markov chains

Definition

A Markov chain is a tuple $(\mathrm{S}, \mathrm{s}_0, \mathrm{E})$ such that S is a finite set of states, and $\mathrm{E}\colon \mathrm{S} \times \mathrm{S} \to [0, 1]$, and for each s we have $\sum_{\mathrm{t} \in \mathrm{S}} \mathrm{E}(\mathrm{s}, \mathrm{t}) = 1$.

- Labeled Markov chains.
- Generate probability space on S^ω.
- Matrix representation.

Theorem

Let M be a Markov chain.

- Probability of reaching some BSCC is 1.
- If M is strongly connected, then for each state t, the probability of reaching t infinitely often is 1.
- Reachability with non-zero (> 0) and almost-sure ($= 1$) probability reduces to graph problems.

Quantitative reachability

Theorem

Let M be a Markov chain and T be a subset of states.

- The probability of reaching T is given by the least solution to $\vec{x} = A\vec{x} + \vec{b}$.
- If we remove states from which T is unreachable, the solution is unique.

Example of non-uniqueness.

Corollary

Let M be a Markov chain and T be a subset of states. We can compute in polynomial time probabilities $p_{s,T}$ of reaching T from s.

Quantitative model checking

Theorem (Model checking ω-regular properties)

Let $\mathrm{M} = (\mathrm{S}, \mathrm{s}_0, \mathrm{E})$ be a Markov chain and let $\mathcal{L} \subseteq \mathrm{S}^\omega$ be an ω-regular languages given by a deterministic Rabin automaton $\mathcal{A}$. We can compute the probability of $\mathbb{L}$ in M in polynomial time in $\mathcal{A}$.

Markov Decision Processes

Definition

A Markov Decision Process (MDP) is a tuple $(\mathrm{S}_1, \mathrm{S}_\mathrm{P}, \mathrm{s}_0, \mathrm{E})$ such that $\mathrm{S}_1, \mathrm{S}_\mathrm{P}$ are a finite set of states, and $\mathrm{E}\colon (\mathrm{S}_1 \cup \mathrm{S}_\mathrm{P}) \times (\mathrm{S}_1 \cup \mathrm{S}_\mathrm{P}) \cup \to [0, 1]$, and for each $\mathrm{s} \in \mathrm{S}_\mathrm{P}$ we have $\sum_{\mathrm{t} \in (\mathrm{S}_1 \cup \mathrm{S}_\mathrm{P})} \mathrm{E}(\mathrm{s}, \mathrm{t}) = 1$.

Markov Decision Processes

Definition

A Markov Decision Process (MDP) is a tuple (S_1, S_P, s_0, E) such that S_1, S_P are a finite set of states, and $E\colon (S_1 \cup S_P) \times (S_1 \cup S_P) \cup \to [0, 1]$, and for each $s \in S_P$ we have $\sum_{t \in (S_1 \cup S_P)} E(s, t) = 1$.

Theorem

Let M be an MDP.

- Reachability with non-zero (> 0) and almost-sure ($= 1$) probability reduces to graph problems.
- Quantitative reachability reduces to linear programming.