Robust Monte Carlo Methods for Sequential Planning and Decision Making

Sue Zheng, Jason Pacheco, & John Fisher

Sensing, Learning, & Inference Group
Computer Science & Artificial Intelligence Laboratory
Massachusetts Institute of Technology

November 30, 2017
Diversion Detection

Given sensor characteristics, observations, and ideal network, infer deviations (e.g., unknown network of material diversion).

[Image: http://isis-online.org/section-3]
Announce site inspections to learn causal network structure. Modify sensor configurations to reduce uncertainty of inferences.
Sequential Experiment Design

Experimental choice is a form of planning.

[Image: Liepe et al., 2013]
Develop algorithm for **sequential Bayesian inference** and **planning** with the following desiderata,

- **Information theoretic** approach to planning
- Effective even for complex models
- Provable theoretical guarantees
- Maximally **reuses computation** in inference and planning phases
Probabilistic Model & Inference

Joint probability model: \( p(x, y) = p(x)p(y|x) \)

- **Prior belief in diversion network**
- **Likelihood of observations given structure**

Posterior belief in network structure given observations:

\[
p(x|y) = \frac{p(x)p(y|x)}{p(y)}
\]
Decision-Driven Observations

Configuration variable $d = \{1, \ldots, D\}$ controls observation model. E.g., announcing inspection of site A affects sensor observations.

Choose configuration to **reduce posterior uncertainty**
Entropy

\[ H(X) = \mathbb{E}[- \log p(X)] \]

Encodes uncertainty, more reliable than variance in many cases (multimodality)

**Coin flip example:** \( X \sim \text{Bernoulli}(p) \)

Entropy maximized when coin is fair, greatest uncertainty
Mutual Information

What would be uncertainty of $X$ if we knew $Y$?

$$I(X; Y) = H(X) - H(X|Y)$$

Prior uncertainty $\quad$ Expected posterior uncertainty

I will drop explicit dependence on configuration variable $d$ when clear from context:

$$I(X; Y|d) \Rightarrow I(X; Y)$$
Information Theoretic Planning

- **Execute Plan**
  
  \[ y \sim p(\cdot | x; d) \]

- **Execute:** Draw new observation \( Y = y \)

- **Planning:** Choose most informative decision
  
  \[ d^* = \arg \max_d I(X; Y | d) \]
Information Theoretic Planning

- **Execute**: Draw new observation $Y = y$
- **Inference**: Update posterior belief

\[
p(X|y) \propto p(X)p(y|X)
\]
Information Theoretic Planning

**Execute:** Draw new observation $Y = y$

**Inference:** Update posterior belief

$$ p(X|y) \propto p(X)p(y|X) $$

**Planning:** Choose most informative decision

$$ d^* = \arg\max_d I(X; Y|d) $$
Closed-Loop Greedy Planning

Condition on observed information during planning

\[ d_t^{\text{greedy}} = \arg \max_d I(X; Y_t|y_1^{t-1}; d) \]
Closed-Loop Greedy Planning

Observation sequence:

Decision sequence:

Unknown quantity:

Condition on observed information during planning

\[ d_{t}^{\text{greedy}} = \arg \max_{d} I(X; Y_{t}|y_{t}^{t-1}; d) \]
Closed-Loop Greedy Planning

Condition on observed information during planning

\[ d_t^{\text{greedy}} = \arg \max_d I(X; Y_t | y_1^{t-1}; d) \]
Closed-Loop Greedy Planning

Observation sequence:

Decision sequence:

Condition on observed information during planning

\[ d_t^{\text{greedy}} = \arg \max_d I(X; Y_t | y_1^{t-1}; d) \]
Closed-Loop Greedy Planning

Observation sequence:

Decision sequence:

Condition on observed information during planning

$$d_t^{\text{greedy}} = \arg \max_d I(X; Y_t | y_1^{t-1}; d)$$
Estimating Mutual Information

Mutual information typically lacks closed-form:

\[ I(X; Y) = \mathbb{E} \left[ \log \frac{p(x, y)}{p(x)p(y)} \right] \]

Evidence integrates every latent configuration:

\[ p(y) = \int p(x, y) dx \]

Can use Monte Carlo integration to estimate integrals from joint samples:

\[ \{x_i, y_i\}_{i=1}^N \sim p(x, y) \]

Empirical estimate is sensitive to outliers in small sample regime
Robust Estimation of Information

- Robust M-estimator is solution to root equation
  \[ \sum_i \psi(\alpha(\theta_i - \hat{\theta})) = 0 \text{ where } \theta_i = \log \frac{p(x_i, y_i)}{p(x_i)p(y_i)} \]

- Influence function \( \psi \) reduces impact of outliers

- Provides quality guarantee in finite sample setting

[Images: Catoni, 2010]
At time $t$ execute plan and draw observation:

$$y_t \sim p(y|x; d_t^{\text{greedy}})$$

Do posterior inference via MCMC samples

$$\{x_i\}_{i=1}^N \sim p(x|y_1^t)$$
For each decision draw samples:
\[
\{x_i\}_{i=1}^N \sim p(x|Y_{t+1}, y_t; d)
\]

Robust estimation of model evidence:
\[
\hat{p}(Y_{t+1}|y_t; d)
\]

Can reuse MCMC samples with importance sampling
For each decision robust MI estimate:

$$\hat{I}(X; Y_{t+1}|y_t^1; d)$$

Greedy planning:

$$d_{t+1}^{\text{greedy}} = \arg \max_d \hat{I}(X; Y_{t+1}|y_t^1; d)$$
Sequential Inference & Planning

- MCMC
  - Run MCMC only when necessary

- Robust Evidence Estimate
  - Avoid additional samples during planning

- Robust MI Estimation

Significantly reduces computation in inference and planning stages through sample reuse.

Algorithmic details are similar to a particle filter...
Asymptotic Results

Theoretical bounds on estimators ensure high quality decisions.

\[ \sqrt{N}(\hat{H} - H) \overset{d}{\to} \mathcal{N}(0, \cdot) \]

Establish central limit theorem as number of samples \( N \to \infty \).

Show that estimators are consistent and approximately Normal with variance \( \Theta\left(\frac{1}{N}\right) \).
Finite Sample Bounds

In practice, finite sample bounds are preferred over asymptotic results.

\[ H(Y) + b - \text{const} < \hat{H}_Y < H(Y) + b + \text{const} \quad \text{w.p.} \geq 1 - 2\epsilon \]

where \( b = \text{KL}(p(Y) || \hat{p}(Y)) \)

Estimates are biased but deviation is bounded w.h.p. through use of robust estimator.

[Image: Catoni, 2010]
Diversions in Nuclear Fuel Cycle

Identify sites for inspection announcement.
Performing an **intervention** to learn causal network structure.
Causal Network Inference

Nodes interact linearly according to directed acyclic graph (DAG) structure.

Interaction Weight: $w$

Directed Acyclic Graph: $G$

Node Observation: $X$
Causal Network Inference

Graph structure and interaction weights unknown.
Causal Network Inference

Covariance Matrix

Possible Graphs

Cannot determine causality from correlations, need to perform active interventions

 Clamp node to fixed value.
Causal Network Inference

Perform optimal intervention

Evalutate candidate interventions

Before

After (inferred)

Intervention 1 Intervention 2 Intervention k

Expected change in belief

Initial data

Calculate/update belief

Probability

Candidate network

[Image: Cho et al., 2016]
**Causal Network Inference**

**Model:**

\[
x_j | x_{\text{Pa}(j)}, w_j, G \sim \mathcal{N}(w_j^T x_{\text{Pa}(j)}, \sigma_j^2)
\]

\[
w_j | G \sim \mathcal{N}(\cdot)
\]

\[G \sim \text{Uniform-DAG}\]

**Planning:**

Previous \(t - 1\) observations:

\[\mathcal{X} = \{x(1), \ldots, x(t-1)\}\]

Select intervention to maximize mutual information:

\[d^* = \arg \max_{d} I(G; \hat{X} | \mathcal{X}, d)\]

Clamp node \(x_{d^*} = 0\) and observe remaining nodes.
Robust planning selects most informative interventions in early iterations.
Causal Network Inference

- Median (solid) best/worst (dashed) out of 50 runs
- More informative experiments than Random
- In early iterations chooses more informative experiments compared to Cho et al., 2016
Sequential Bayesian inference and planning applies to complex models (more interesting applications...)

Theoretical guarantees on estimator quality for costly decisions.

Scale up to larger problems and bound probability of incorrect selection.
Measurement Selection

- Sensor Types
  - Satellite
  - Flyby
  - Earth-based

- Sensing Modes
  - EO/SAR/IR
  - Hyperspectral
  - Radio Freq

- Measurement Set

- Time

- Unknown Quantity

- Selected Measurement

- Optimal Selection

Consortium for Verification Technology
Plan Execution Costs

$$d^* = \arg \max_d I(X; Y|d) - \lambda R(d)$$

- Plan is feasible if information justifies cost:
  $$I(X; Y|d) > \lambda R(d)$$

- Possible costs for this application:
  - Sensor costs
  - Power consumption
Monte Carlo Integration

Need to compute expected values of the form:

$$\mathbb{E}[f(x)] = \int p(x)f(x)\,dx$$

Draw samples from the distribution:

$$\{x^i\}_{i=1}^N \sim p(x)$$

Monte Carlo integration is empirical mean

$$\mathbb{E}[f(x)] \approx \frac{1}{N} \sum_{i=1}^N f(x^i)$$

Good statistical properties as $N \to \infty$, but problematic for finite samples
Asymptotic and Finite Sample Bounds

Asymptotic Bounds:

\[ \sqrt{N} \hat{H}_Y \xrightarrow{d} \mathcal{N} \left( H(Y), \mathbb{E}_y \left[ \frac{\sigma_x^2 (p(Y|X))}{M p^2(Y)} + \sigma^2 (\log p(Y)) \right] \right) \]

\[ \sqrt{N} \hat{H}_Y|_X \xrightarrow{d} \mathcal{N} \left( H(Y|X), \sigma^2 (\log p(Y|X)) \right) \]

Finite-Sample Bounds:

Assuming \( N > 2(1 + \log \epsilon^{-1}) \)

\[ H(Y) + b - c < \hat{H}_Y < H(Y) + b + c \quad \text{w.p.} \quad \geq 1 - 2\epsilon \]

\[ b = \mathbb{E} \left[ \log \frac{p(Y)}{\hat{p}(Y; X)} \right] \]

\[ c = \frac{1 + \log \epsilon^{-1}}{1 - (1 + \log \epsilon^{-1})/N} \sqrt{\frac{\sigma^2}{2N}} \]
Robust is more consistent in intervention selection.

Average realized gain at $t = 1$ indicates intervention at node 6 is optimal.