Fast Fourier Transform Reductions for Bayesian Network Inference

Vincent Hsiao¹, Dana Nau¹, Rina Dechter²

University of Maryland, College Park¹, University of California, Irvine²

Background

Bayesian Network (BN)
- Bayesian Network: graphical model (X,D,F)
  - Variables: \( X = \{X_1, X_2, ..., X_N\} \)
  - Domains: \( D = \{D_{X_1}, D_{X_2}, ..., D_{X_N}\} \)
  - Parent Functions: \( F = \{F_1, F_2, ..., F_N\} \)

Causal Independence (CI)
- Probabilistic relationship between a set of causes \( \{c_1, ..., c_m\} \) and an effect \( e \), such that:
  \[
  \begin{align*}
  e &:= b_1 \cdot b_2 \cdot ... \cdot b_n \\
  \text{where each hidden variable } b_i \text{ is some probabilistic } \\
  \text{function of its corresponding } c_i \text{ and } * \text{ is some} \\
  \text{commutative and associative binary operator}
  \end{align*}
  \]
- Network Transformation: Given a CI fragment \((X,D,F)\) with a set of causes \( \{c_1, ..., c_m\} \), an effect \( e \), and hidden variables \( \{b_1, ..., b_n\} \), a network transformation is a new network \((X',D',F')\) constructed over some computational ordering:
  \[
  e = c_1 \cdot b_2 \cdot ... \cdot b_n \\
  \text{with new intermediate variables:}
  g_i := b_i \cdot g_{i-1}
  \]
- Algorithms such as ci-elim-bel (Zhang and Poole, 1996; Rish and Dechter, 1998) exploit network transformations to accelerate bucket elimination

Applications
- Distributed Computing, Fault Tree Analysis
- N different resource providers with stochastic availability (k-out-of-n model)
  

Problem Statement

Computing Random Variable Sums
- Given a source-sink network:
  \[
  \begin{align*}
  &\text{source-sink network (2)} \\
  &\text{The size of the parent function for the node } E \text{ expressed} \\
  &\text{as a conditional probability table (CPT) is exponential in} \\
  &\text{the number of } Y \text{ nodes}
  \end{align*}
  \]
- We would like to reduce the size of the CPT of the network. This is useful in applications where one wants to find the marginal distribution of node E as in:
  - Test score prediction
  - Evolutionary games

From the convolution theorem and the application of the FFT, we know that:

**Theorem 2.4 (FFT Time Complexity).** For i.i.d random variables \( X = \{X_1, X_2, ..., X_n\} \) where each variable has a domain size \( |D| \) and their sum \( Z = \sum X_i \), computing \( P(Z = k) \) using the FFT will take time \( O(N|D| \log(N|D|)) \). For non-i.i.d variables, the time complexity is \( O(N^2|D| \log(N|D|)) \).

It follows that:

**Theorem 3.1 (FFT Reduction).** Let \( B = \{X,D,F\} \) with \( X = \{S,Y_1,...,Y_N,E\} \) be a source-sink network with N i.i.d. paths. The network can be transformed into \( \{X',D',F'\} \) such that \( X' = \{S,E\} \) reducing the CPT for \( E \) from size \( O(|D|_X |D|_E) \) to size \( O(|D|_X |D|_E) \) in \( O(|D|_X \log R) \) time where \( R \) is the numerical range of random variable \( E \).

Furthermore, it can be shown that certain computations performed in the algorithm ci-elim-bel can also be formulated as convolutions, leading to an improved algorithm: ci-elim-FFT

FFT Reduction

Experimental Setup

**Experimental Setup (Set 1)**
- Evaluate FFT reduction for inference on a selection of networks with summation-CI
- Compare with:
  - Vanilla Bucket Elimination (Naive)
  - Temporal Decomposition (Temporal)
- Test inference on three types of networks (1), (2), (3)
  - Supply demand network (3)

**Experimental Setup (Set 2)**
- Evaluate ci-elim-FFT compared to ci-elim-bel on general two layer additive networks (4)

Future Research

- Explore integration into existing lifted variable elimination algorithms

Acknowledgements

This work supported in part by NSF grant IIS-2008516 and AFOSR grant 19105W035.