Smoothing Structured Decomposable Circuits

Andy Shih\textsuperscript{1}  Guy Van den Broeck\textsuperscript{2}  Paul Beame\textsuperscript{3}  Antoine Amarilli\textsuperscript{4}

\textsuperscript{1}Stanford University  
\textsuperscript{2}University of California, Los Angeles  
\textsuperscript{3}University of Washington  
\textsuperscript{4}LTCI, Télécom Paris, IP Paris

Dec 2019
Probabilistic Circuits

Tractable computation graph, encoding a distribution.

SOTA for:

- Inference algorithms for PGMs
- Inference algorithms for probabilistic programs
- Discrete density estimation

Check out:

**Tractable Probabilistic Models**: (UAI19 / AAAI20 tutorial)
Tractability

Different combination of properties leads to different families of circuits

<table>
<thead>
<tr>
<th></th>
<th>SPN</th>
<th>AC</th>
<th>PSDD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decomposability</td>
<td>✓</td>
<td>✓</td>
<td>✓(S)</td>
</tr>
<tr>
<td>Determinsim</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Smoothness</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Pr(evid)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Marginal</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>MPE</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Marginal MAP</td>
<td>✗</td>
<td>✗</td>
<td>✓*</td>
</tr>
<tr>
<td>Expectation</td>
<td>✗</td>
<td>✗</td>
<td>✓*</td>
</tr>
</tbody>
</table>

...with different tractability properties.
Smoothness

Definition
A circuit is **smooth** if for every pair of children $c_1$ and $c_2$ of a $\oplus$-gate, $\text{vars}_{c_1} = \text{vars}_{c_2}$.

**Figure:** Two equivalent circuits computing $(x_0 \otimes x_1) \oplus x_2$. The left one is not smooth and the right one is smooth.
Smoothing a Circuit: Naive Quadratic Algorithm

- Go to each gate $O(m)$ and fill in each variable $O(n)$
- Complexity $O(nm)$
- Problematic when $n \geq 1,000$ and $m \geq 1,000,000$

Our near-linear smoothing algorithm: $O(m \cdot \alpha(m, n))$
Smoothing a Circuit: Missing Intervals

Key Insight: missing variables for each gate form two intervals (in the inorder traversal of vtree).

We need to fill in $2m$ intervals.

**Figure:** $A \setminus B$ forms two intervals
Semigroup Range Sum

Theorem

Given $n$ variables defined over a semigroup and $m$ intervals, the sum of all intervals can be computed using $O(m \cdot \alpha(m, n))$ additions [Chazelle and Rosenberg 1989].

$\alpha(m, n)$ is the inverse Ackermann function, which grows very slowly.

*The original theorem only bounds the number of additions. We bound the number of total operations.*
Takeaways

- Probabilistic circuits can encode complex distributions.
- They can compute exact likelihoods, marginals, and more
  - But only if they are smooth.
- Best smoothing algorithm was quadratic.
- We propose a near linear time smoothing algorithm.
Thanks!

Poster: East Exhibition Hall B+C #182, 10:45AM

Code: https://github.com/AndyShih12/SSDC

Contact: andyshih@cs.stanford.edu