Smoothing Structured Decomposable Circuits

Andy Shih[1] Guy Van den Broeck[2] Paul Beame[3] Antoine Amarilli[4]

[1]Stanford University
[2]University of California, Los Angeles
[3]University of Washington
[4]LTCI, Télécom Paris, IP Paris

NeurIPS 2019

Probabilistic Circuits

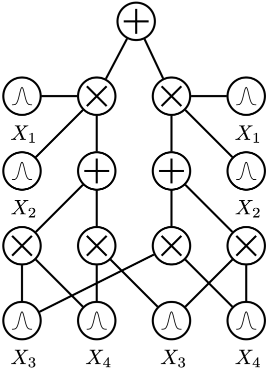
Tractable computation graph, encoding a distribution.

SOTA for:

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

Exact likelihoods and partition function!

Gaining popularity:
Tractable Probabilistic Models: (UAI19 / AAAI20 tutorial)



Tractability

Different combination of properties leads to different families of circuits

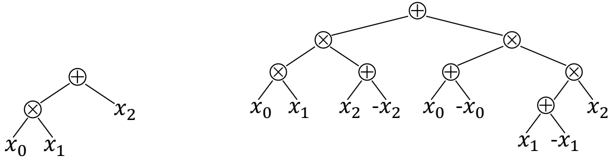
	SPN	AC	PSDD
Decomposability	✓	✓	✓(S)
Determinism	✗	✓	✓
Smoothness	✓	✓	✓
Pr(evid)	✓	✓	✓
Marginal	✓	✓	✓
MPE	✗	✓	✓
Marginal MAP	✗	✗	✓*
Expectation	✗	✗	✓*

...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, $vars_{c_1} = vars_{c_2}$.



(a) A circuit. **(b)** A smooth circuit.

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: Prior Work

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

Our near-linear smoothing algorithm: $\boldsymbol{O(m \cdot \alpha(m, n))}$

Smoothing a Circuit: Our Work

Key Insight: missing variables for each gate form two intervals.

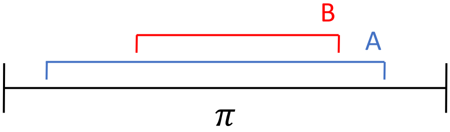


Figure: $A \setminus B$ forms two intervals

We need to fill in $2m$ 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 computations.

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