Learning Probabilistic Sentential Decision Diagrams Under Logic Constraints by Sampling and Averaging

Renato Lui Geh, Denis Deratani Mauá
Department of Computer Science, Institute of Mathematics and Statistics, University of São Paulo, Brazil
{renatolg,ddm}@ime.usp.br

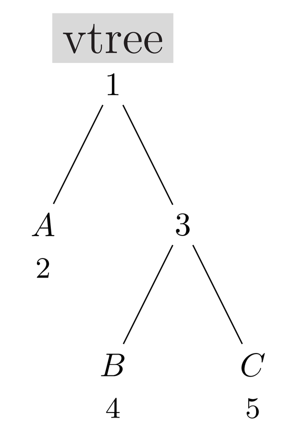
1. Motivation

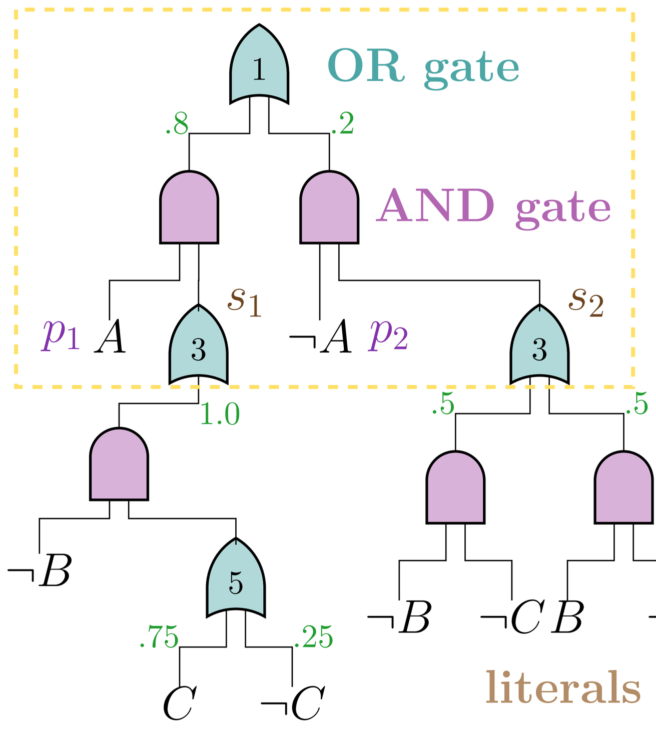
Probabilistic Sentential Decision Diagrams (PSDDs):

- Structured Decomposable probabilistic circuits
- Encode certain knowledge as logic constraints
- Encode uncertain knowledge as probabilities
- Interpretable syntax
- Many inferences are exact and tractable:
  - ☑ Evidence
  - ☑ Marginals
  - ☑ MLE Parameter Learning
  - ☑ Most Probable Explanation
  - ☑ Expectations
  - ☑ KL-divergence

A	B	C	Pr
0	0	0	**0.1**
0	1	0	**0.1**
1	0	0	**0.2**
1	0	1	**0.6**

s.t. $(A \to \neg B) \wedge (C \to A)$





- PSDD circuit represents recursive decomposition of formula:

$$\bigvee_{i=1}^{k} (p_i \wedge s_i),$$ where each *prime* p_i and *sub* s_i are logical formulae

Existing PSDD learners:

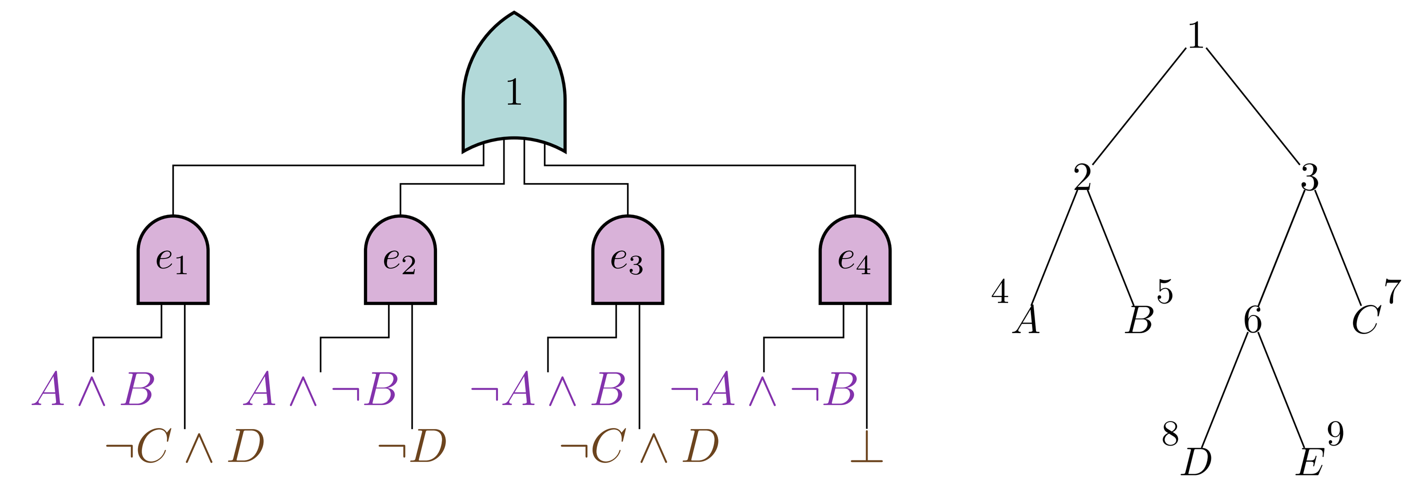
- require an initial PSDD encoding the support;
- scale poorly to complex formulae and/or high dimension.

This Work: How to effectively learn PSDDs s.t. complex formula?

2. SamplePSDD

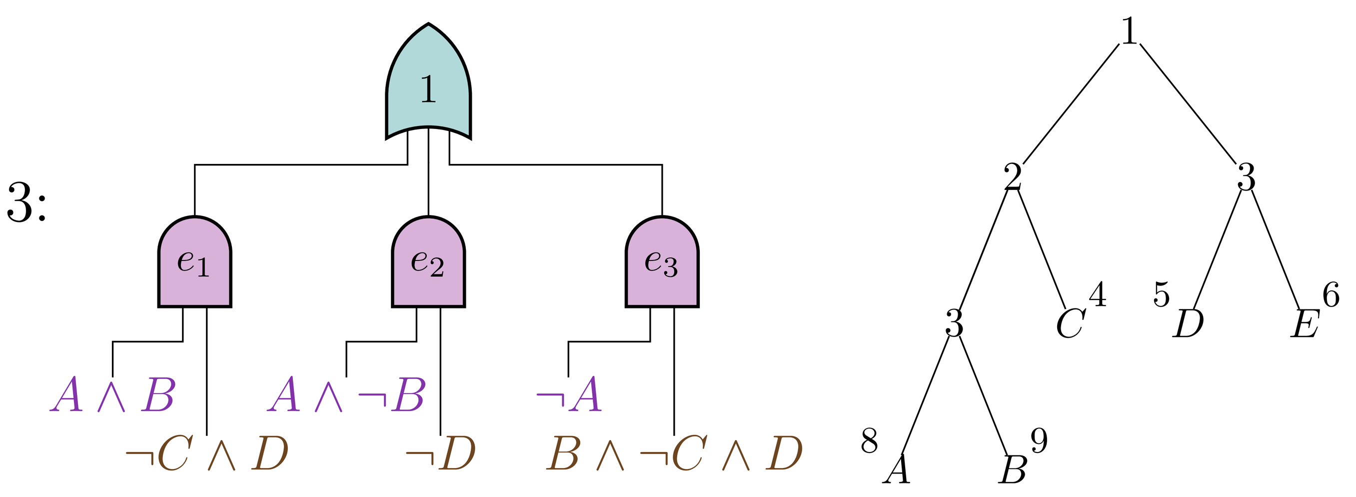
- Common assumption: primes p_i are conjunctions of literals.

$$\phi(A,B,C,D) = (A \wedge \neg B \wedge \neg D) \vee (B \wedge \neg C \wedge D)$$

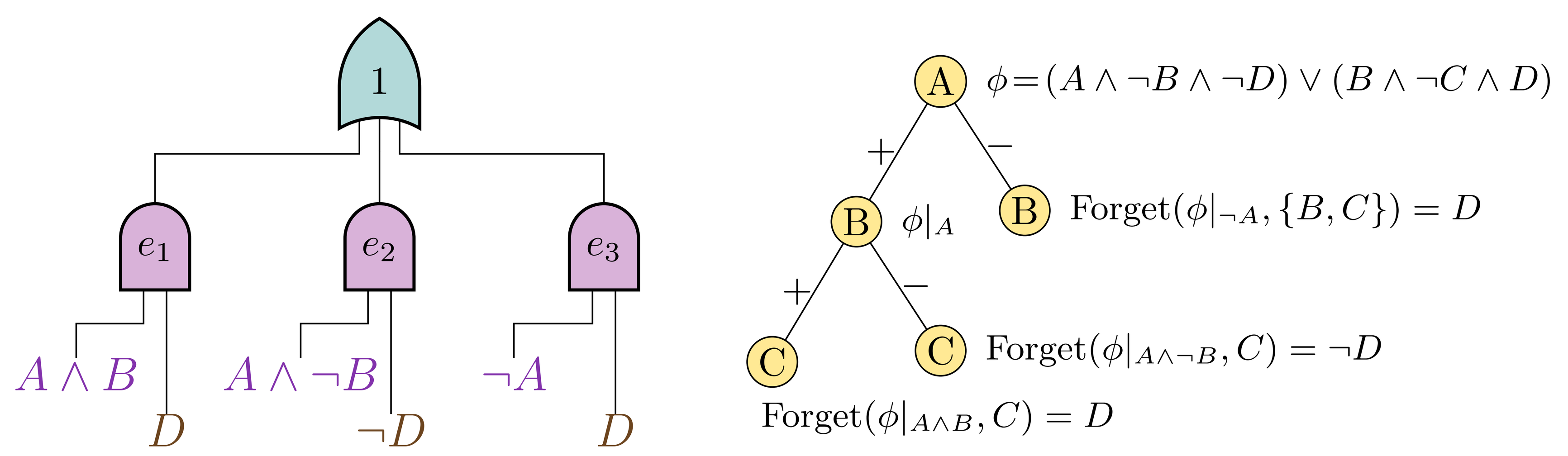


- **Problem:** Size of circuit is exponential in the size of p_i
- **Solution:** randomly sample a bounded number (k) of p_i
- **But:** this violates structure decomposability

Example, $k = 3$:



- **New solution:** Relax logical constraints ϕ



3. Experiments

Evaluation: we sample 30 PSDDs and use 5 ensemble strategies:

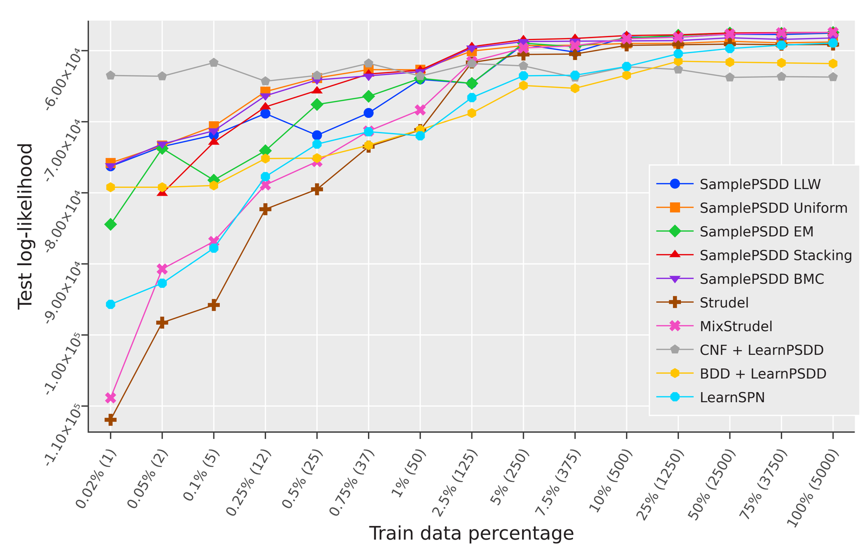
1. Likelihood weighting (LLW),
2. Uniform weights,
3. Expectation Maximization (EM),
4. Stacking,
5. Bayesian Model Combination;

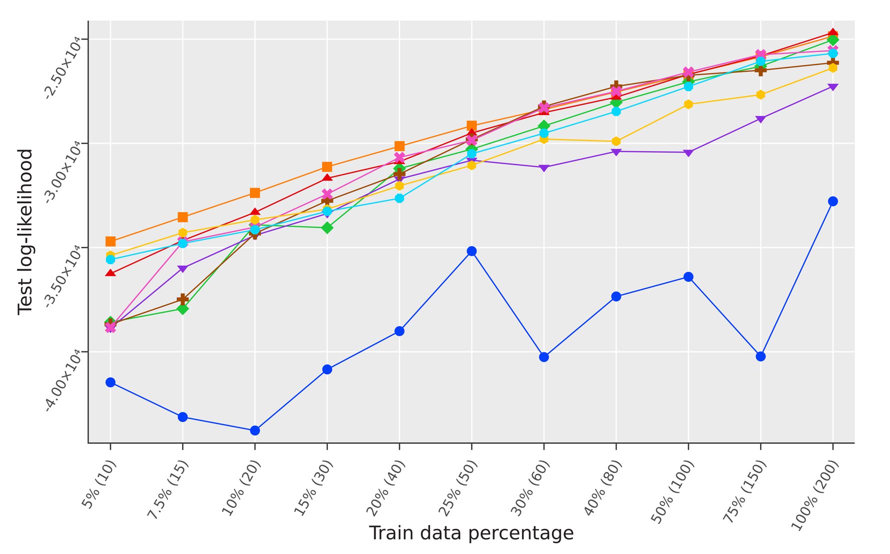
comparing with Strudel, LearnPSDD and LearnSPN.

Datasets: we evaluate with 5 data + knowledge as logic constraints:

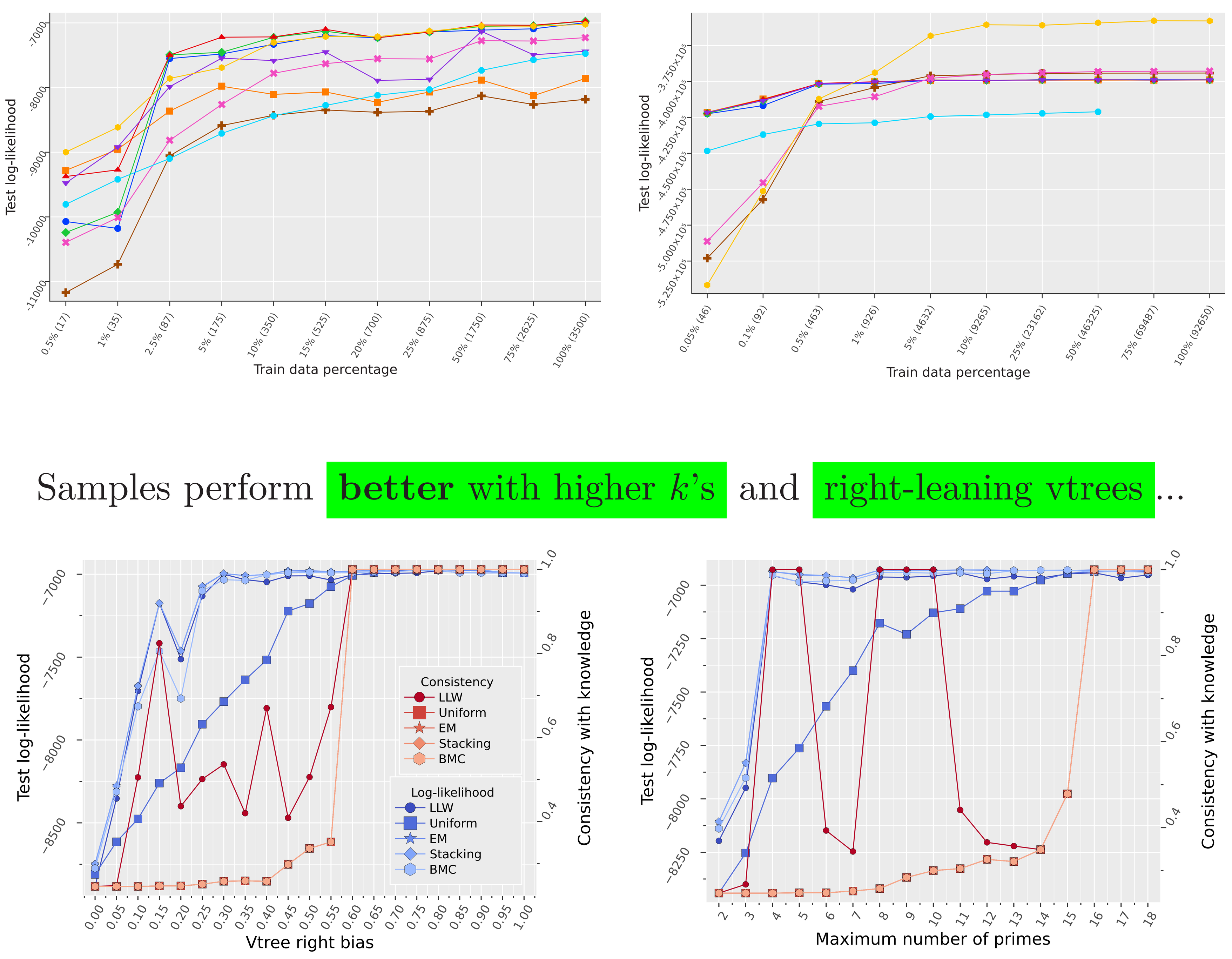
Dataset	#vars	#train	ϕ's size
LED	14	5000	23
LED + Images	157	700	39899
Sushi Ranking	100	3500	17413
Sushi Top 10	10	3500	37
Dota 2 Games	227	92650	1308

Our approach fares **better with fewer** data, yet remains **competitive** under **lots of data**.





Samples perform **better with higher k's** and **right-leaning vtrees**...



...but at a **cost** to complexity.