Efficient Search-Based Weighted Model Integration

Zhe Zeng, Guy Van den Broeck

{zhezeng, guyvdb}@cs.ucla.edu

UCLA

PROBLEM DEFINITION

- Satisfiability Modulo Theories (SMT) generalize SAT
- SMT(LRA) : Boolean combination of atomic propositions over Boolean variables (e.g. A, B), and of atomic linear real arithmetic (LRA) formulas over real variables (e.g. x < y + 5)
- Weighted Model Integration (WMI) over SMT(LRA):

$$\mathrm{WMI}(\theta, w \mid \boldsymbol{x}, \boldsymbol{b}) = \sum_{\boldsymbol{b}^\star \in \mathbb{B}^m} \int_{\theta(\boldsymbol{x}, \boldsymbol{b}^\star)} w(\boldsymbol{x}, \boldsymbol{b}^\star)\, d\boldsymbol{x}.$$

- Example: house price model

$$\gamma_i = \begin{cases} (price_i < 10 \cdot sqft_i + 1000) \lor (price_i < 20 \cdot sqft_i + 100) \\ (0 < price_i < 3000) \land (0 < sqft_i < 200) \end{cases}$$

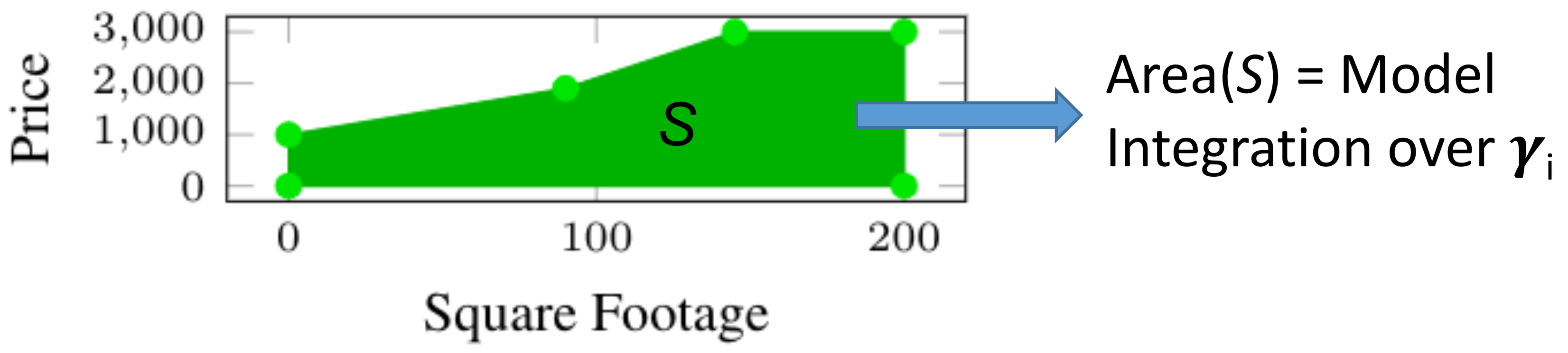


Figure 1. Feasible region of SMT(LRA) model $\boldsymbol{\gamma}_i$.

MOTIVATION

- **Goal**: to develop efficient inference algorithm for WMI problems on hybrid domains
- To leverage **context-specific independence** in model integration (MI) problem to speed up inference

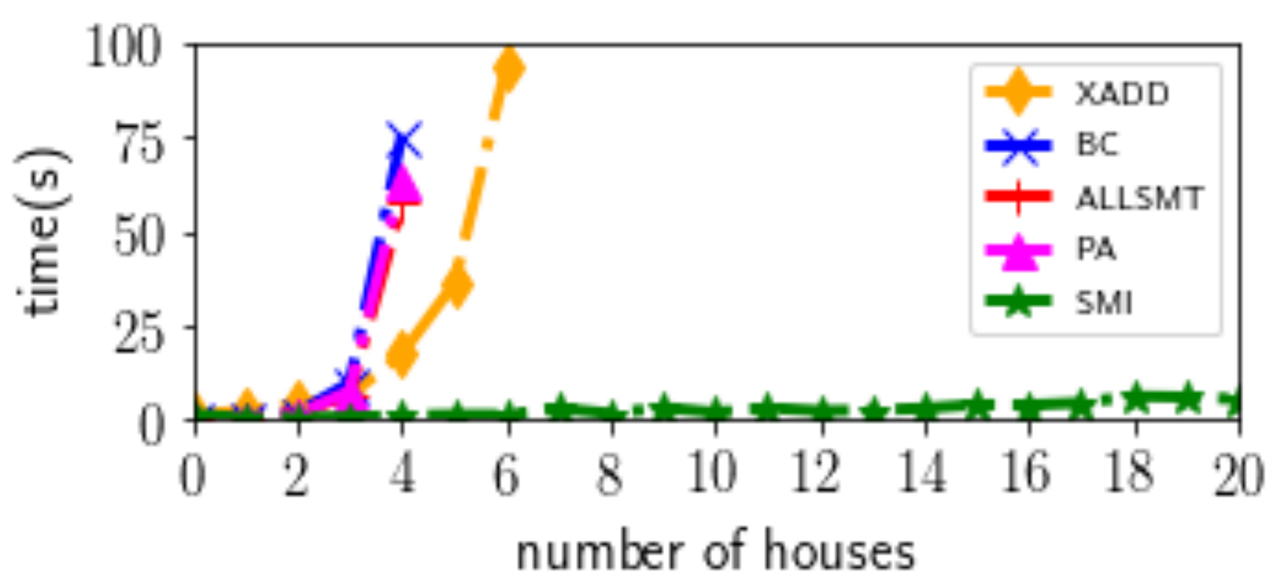


Figure 2. Model integration runtime on independent house price model.

MODEL INTEGRATION IS ALL U NEED

- Boolean to real:

$$\begin{bmatrix} b & w(b) = 2 \\ \neg b & w(\neg b) = 3 \end{bmatrix}$$

$$\begin{bmatrix} \lambda_b > 0 & w(\lambda_b > 0) = 2 \\ \lambda_b < 0 & w(\lambda_b < 0) = 3 \\ -1 \leq \lambda_b \leq 1 & \end{bmatrix}$$

- weighted to unweighted:

$$\begin{bmatrix} x + y \geq 1 & w(x + y \geq 1) = x^2 y \\ \neg(x + y \geq 1) & w(\neg(x + y \geq 1)) = 1 \end{bmatrix}$$

$$\begin{bmatrix} x + y \geq 1 & \Rightarrow \begin{cases} \bigvee_{i=1,2} 0 \leq p_{x,i} \leq x \\ \leq p_{y,1} \leq y \end{cases} \\ \neg(x + y \geq 1) & \Rightarrow \begin{cases} \bigvee_{i=1,2} 0 \leq p_{x,i} \leq 1 \\ \leq p_{y,1} \leq 1 \end{cases} \end{bmatrix}$$

✓ Focus on MI without loss of generality

GRAPH ABSTRACTION OF SMT

- *Definition (Primal Graph)*: vertices being all variables; edges connecting any two variables in the same clause.
  - Boolean formula:

$$(y \lor x_1) \land (y \lor x_2)$$

  - SMT formula:

$$\begin{cases} (-1 \leq y \leq 1) \\ (-0.5 \leq x_1, x_2 \leq 0.5) \\ (x_1 + 1 \leq y) \lor (y \leq x_1 - 1) \\ (x_2 + 1 \leq y) \lor (y \leq x_2 - 1) \end{cases}$$

Primal Graph

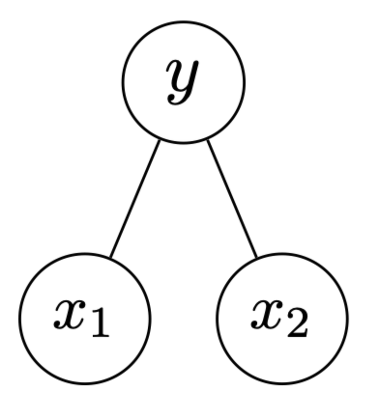


Figure 3. Example primal graph for Both formulas.

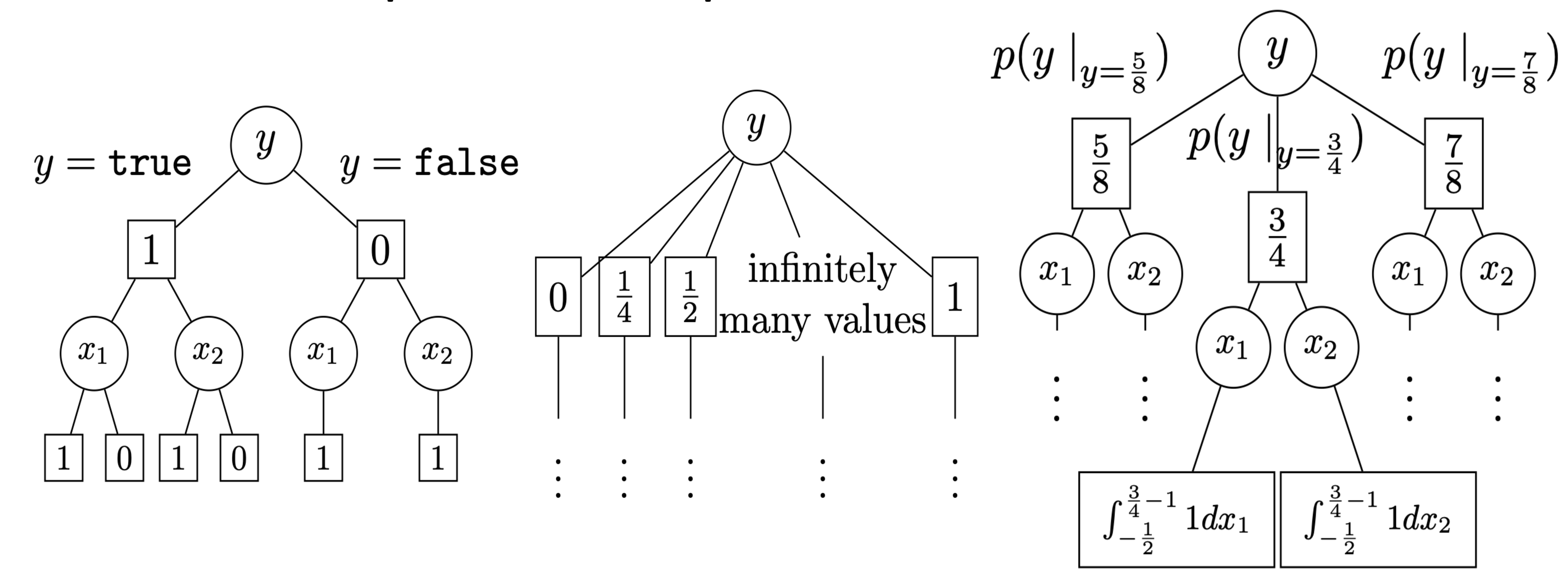
SEARCH-BASED MI

- *Proposition*: MI of SMT(LRA) theory is integration over a univariate piecewise polynomial.

$$MI(\theta \mid \boldsymbol{x}, \boldsymbol{b}) = \sum_{[l,u] \in I} \int_l^u p_{l,u}(y) dy.$$

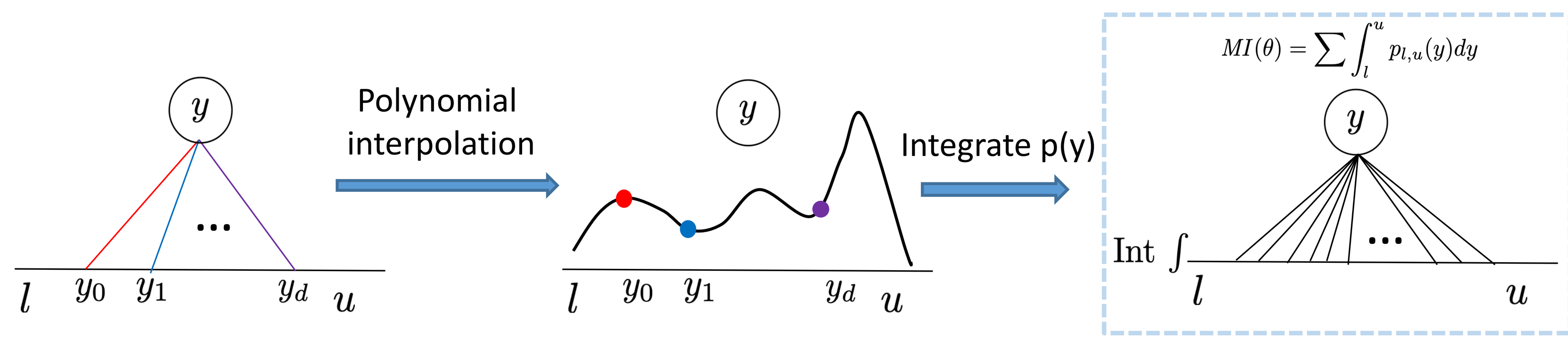
where *I* are intervals over which *p(y)* is polynomial.

- This observation allows search space leveraging context-specific independence



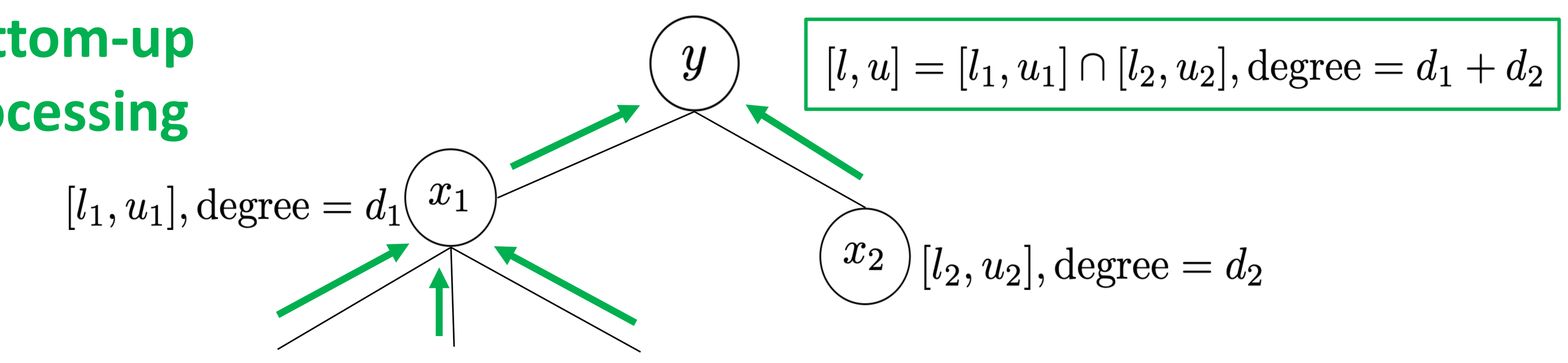
(a) Discrete And/Or search tree with Boolean variables (b) Infinite search tree with real variable y (c) Our proposed finite search tree on interval [l, u]

- From finite branches (instantiations) to integration



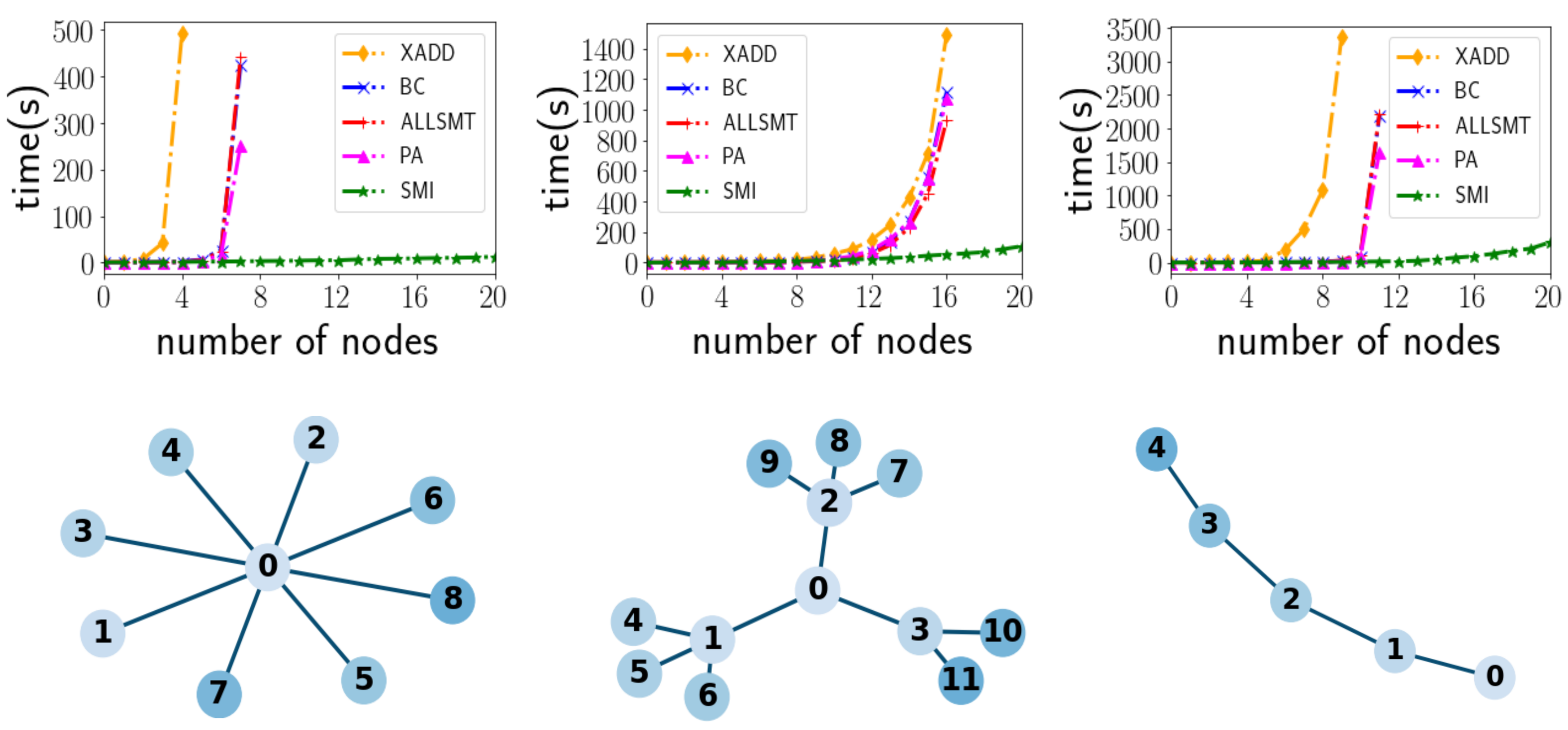
- How to find polynomial intervals and degrees?

Bottom-up processing

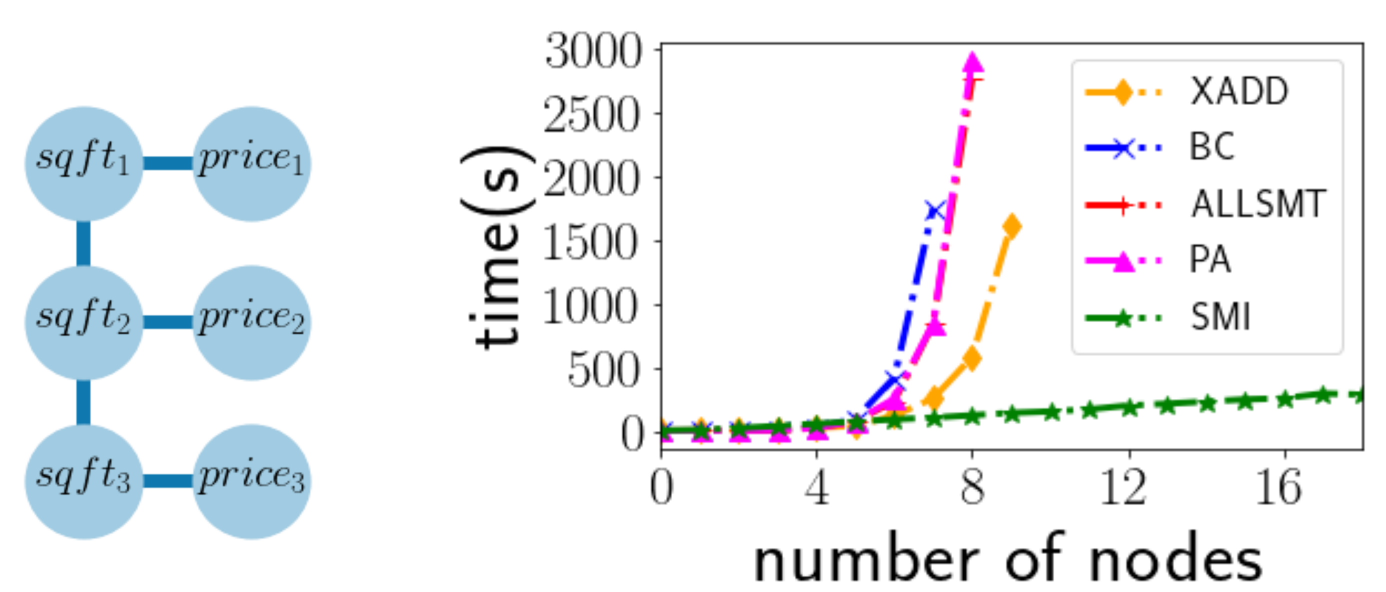


- Complexity analysis: search space size $O(l \cdot (n^3 \cdot c^{h_p})^{h_t})$ is bounded exponentially by the tree height of the primal graph h_p and that of pseudo tree h_t

EXPERIMENTS



(a) Model integration runtime on star, full three-ary tree, and path primal graphs.



(b) Model integration runtime on non-independent house price model.

Acknowledgements: This work is partially supported by NSF grants #IIS-1657613, #IIS-1633857, #CCF-1837129, DARPA XAI grant #N66001-17-2-4032, NEC Research, and gifts from Intel and Facebook Research.