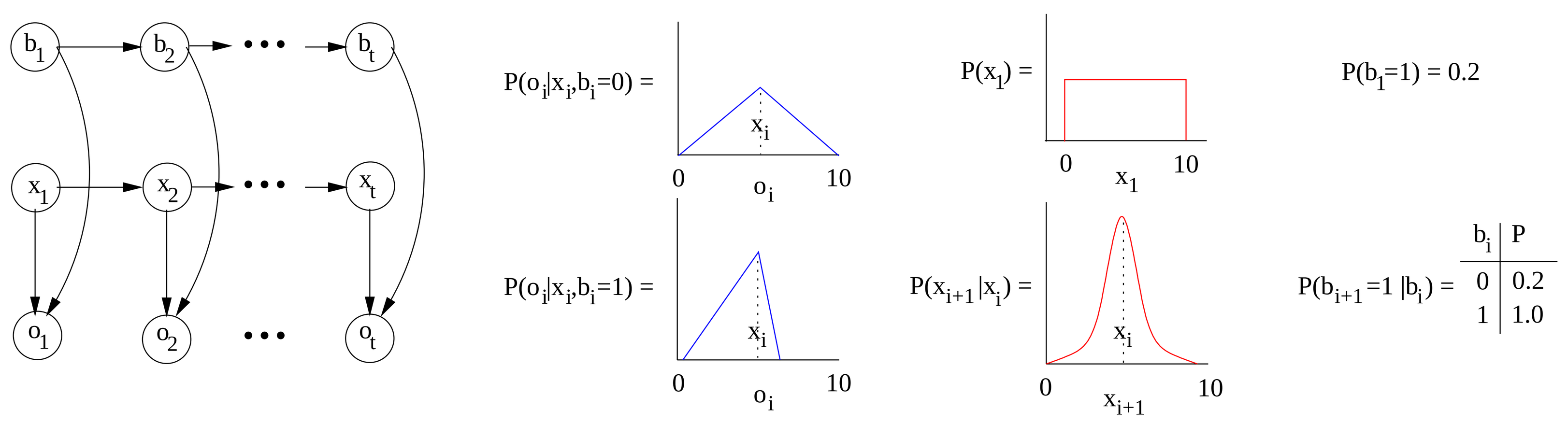
Symbolic Variable Elimination for Discrete and Continuous Graphical Models

Scott Sanner (NICTA & Australian National University), Ehsan Abbasnejad (Australian National University & NICTA)

Highlight

Question: How to do **closed-form exact inference** in discrete and continuous variable (hybrid) **graphical models** with **complex continuous distributions**? E.g.,



Proposal: Represent **factors** by **linear piecewise polynomials**.

Solution: Use **symbolic variable elimination** and **extended ADDs (XADDs)**.

Inference in Hybrid Graphical Models

Compile: **Bayes nets** and **MRFs** $\longrightarrow$ **Factor Graph** with $\mathbf{b} \in \{0,1\}^m$ **and** $\mathbf{x} \in \mathbb{R}^n$:

$$p(\mathbf{b},\mathbf{x}) = \frac{1}{Z}\prod_{f \in F} \Psi_f(\mathbf{b}_f, \mathbf{x}_f)$$

where each $\Psi_f(\mathbf{b}_f, \mathbf{x}_f)$ is a **linear piecewise polynomial case statement**, e.g.,

$$\Psi_f(b_1, x_1, x_2) = \begin{cases} b_1 \wedge x_1 > x_2 : & x_1^2 + x_1 x_2 \\ \neg b_1 \wedge x_1 \le x_2 \wedge x_1 \ge -x_2 : & x_2^3 \\ \neg b_1 \wedge x_1 \le x_2 \wedge x_1 < -x_2 : & x_1^2 x_2^2 \end{cases}$$

(Conditional) Probability Queries: for query $\mathbf{q} = (b_{s+1}, \ldots, b_{s'}, x_{t+1}, \ldots, x_{t'})$ and evidence $\mathbf{e}$ where $\{\mathbf{b} \cup \mathbf{x}\} \setminus \{\mathbf{q} \cup \mathbf{e}\} = (b_1, \ldots, b_s, x_1, \ldots, x_t)$ and $\mathbf{q} = (b_{s+1}, \ldots, b_{s'}, x_{t+1}, \ldots, x_{t'})$:

$$p(\mathbf{q}|\mathbf{e}) = \frac{\sum_{b_1} \cdots \sum_{b_s} \int \cdots \int_{\mathbb{R}^t} \prod_{f \in F} \Psi_f(\mathbf{b}_f, \mathbf{x}_f)\, dx_1 \ldots dx_t}{\sum_{b_1} \cdots \sum_{b_{s'}} \int \cdots \int_{\mathbb{R}^{t'}} \prod_{f \in F} \Psi_f(\mathbf{b}_f, \mathbf{x}_f)\, dx_1 \ldots dx_{t'}}$$

(Conditional) Expectation Queries: for polynomial function $Poly(\mathbf{q})$ over $\mathbf{q}$ (e.g., $Poly(\mathbf{q}) = q_1^2 + q_1 q_2 + q_2^2$) and evidence $\mathbf{e}$, pre-compute $p(\mathbf{q}|\mathbf{e})$ as above, then:

$$\mathbb{E}[Poly(\mathbf{q})|\mathbf{e}] = \int_{\mathbb{R}^{|\mathbf{q}|}} [Poly(\mathbf{q}) \cdot p(\mathbf{q}|\mathbf{e})]\, d\mathbf{q}$$

Contribution: Symbolic Variable Elimination

Algorithm 1: SYMBOLIC VE for conditional probability & expectation queries

```
Input: F, order: a set of factors F, and a variable order for elimination
Output: a set of factors after eliminating each v ∈ order
begin
    foreach v ∈ order do
        f_v ← 1; F_\v ← ∅
        foreach f ∈ F do
            if (f contains v)
            then  f_v ← f_v ⊗ f                 // case multiply: ⊗ (below)
            else  F_\v ← F_\v ∪ {f}
        if (v is boolean)
        then  F ← F_\v ∪ {f_{v=0} ⊕ f_{v=1}}    // case sum: ⊕ (below)
        else  F ← F_\v ∪ {∫_{v=-∞}^{∞} f_v dv}  // case definite integral: ∫_{v=-∞}^{∞} (below)
    return F
```

Case Operations for Symbolic VE

Question: How to do $\oplus$ (similarly $\otimes$) and $\max$ (similarly $\min$) on two cases?

$$\begin{cases} \phi_1 : & f_1 \\ \phi_2 : & f_2 \end{cases} \oplus \begin{cases} \psi_1 : & g_1 \\ \psi_2 : & g_2 \end{cases} = \begin{cases} \phi_1 \wedge \psi_1 : & f_1 + g_1 \\ \phi_1 \wedge \psi_2 : & f_1 + g_2 \\ \phi_2 \wedge \psi_1 : & f_2 + g_1 \\ \phi_2 \wedge \psi_2 : & f_2 + g_2 \end{cases} \qquad \max\left(\begin{cases} \phi_1 : & f_1 \\ \phi_2 : & f_2 \end{cases}, \begin{cases} \psi_1 : & g_1 \\ \psi_2 : & g_2 \end{cases} \right) = \begin{cases} \phi_1 \wedge \psi_1 \wedge f_1 > g_1 : & f_1 \\ \phi_1 \wedge \psi_1 \wedge f_1 \le g_1 : & g_1 \\ \phi_1 \wedge \psi_2 \wedge f_1 > g_2 : & f_1 \\ \phi_1 \wedge \psi_2 \wedge f_1 \le g_2 : & g_2 \\ \vdots : & \vdots \end{cases}$$

Question: How to do $\int_{-\infty}^{\infty} f(x_1)\, dx_1$ for case statement $f(x_1)$? Expand, swap $\sum$, $\int$:

$$\int_{-\infty}^{\infty} f(x_1) dx_1 = \int_{-\infty}^{\infty} \sum_i \mathbb{I}[\phi_i] \cdot f_i \, dx_1 = \sum_i \boxed{\int_{-\infty}^{\infty} \mathbb{I}[\phi_i] \cdot f_i \, dx_1}$$

Hence, we just focus on integrating a single partition: $\boxed{\int_{-\infty}^{\infty} \mathbb{I}[\phi_i] \cdot f_i \, dx_1}$, e.g., let

$$f_i := x_1^2 - x_1 x_2$$
$$\phi_i := [x_1 > -1] \wedge [x_1 > x_2 - 1] \wedge [x_1 \le x_2] \wedge [x_1 \le x_3 + 1] \wedge [x_2 > 0]$$

Derive lower/upper bounds LB/UB from ϕ_i using $\max$/$\min$, integrate, evaluate:

$$LB := \max(x_2 - 1, -1) = \begin{cases} x_2 - 1 > -1 : & x_2 - 1 \\ x_2 - 1 \le -1 : & -1 \end{cases} \qquad UB := \min(x_2, x_3 + 1) = \begin{cases} x_2 < x_3 + 1 : & x_2 \\ x_2 \ge x_3 + 1 : & x_3 + 1 \end{cases}$$

$$\int_{-\infty}^{\infty} \mathbb{I}[\phi_i] \cdot f_i \, dx_1 = \mathbb{I}[x_2 > 0] \int_{x_1 = LB}^{UB} (x_1^2 - x_1 x_2)\, dx_1 = \mathbb{I}[x_2 > 0] \left[\frac{1}{3}x_1^3 - \frac{1}{2}x_1^2 x_2 \Big|_{x_1 = LB}^{x_1 = UB} \right]$$

$$= \mathbb{I}[x_2 > 0] \otimes \left[\left(\frac{1}{3} UB \otimes UB \otimes UB \ominus \frac{1}{2} UB \otimes UB \otimes (x_2) \right) \ominus \left(\frac{1}{3} LB \otimes LB \otimes LB \ominus \frac{1}{2} LB \otimes LB \otimes (x_2) \right) \right]$$

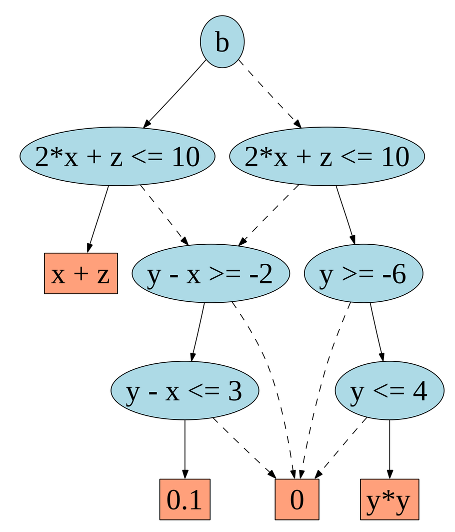
$$\otimes\, \mathbb{I}[x_2 > x_2 - 1] \otimes \mathbb{I}[x_2 > -1] \otimes \mathbb{I}[x_3 + 1 > x_2 - 1] \otimes \mathbb{I}[x_3 + 1 > -1]$$ // **reason for constraints: enforce** $UB > LB$

Extended ADD (XADD)

Question: How to avoid blow-up in case statements during operations?

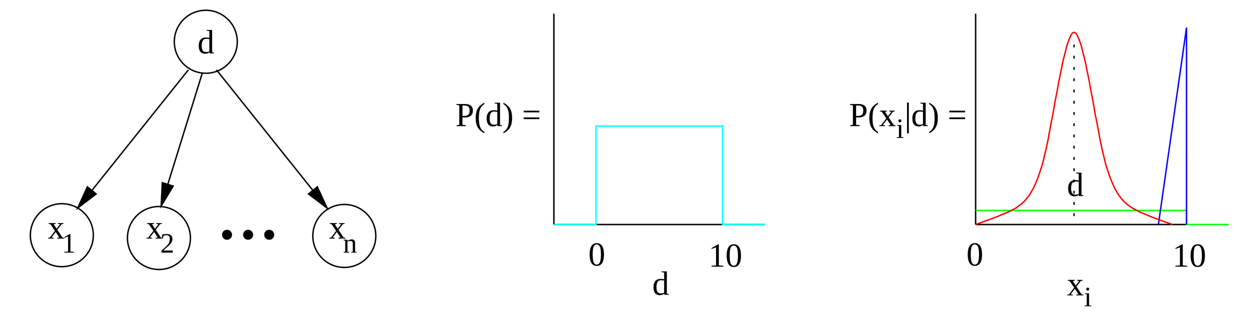
Answer: Extend algebraic decision diagram (ADD) to represent cases $\longrightarrow$ XADD:

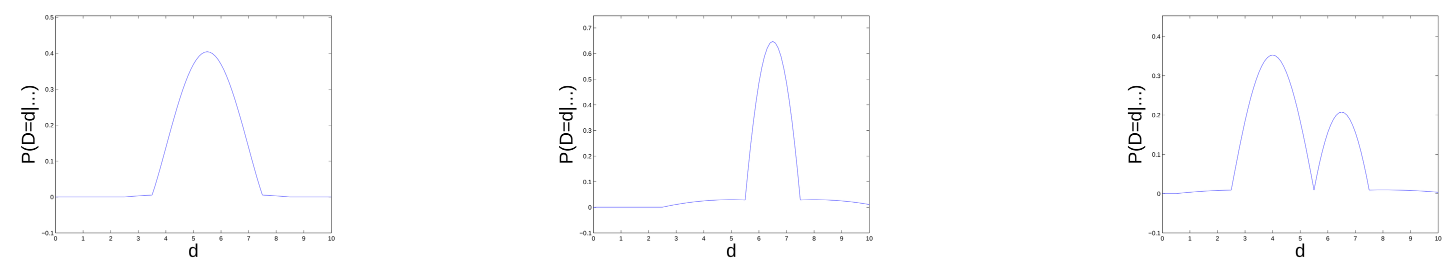
$$\begin{cases} b \wedge 2x + z \le 10 : & x + z \\ b \wedge 2x + z > 10 \wedge y - x \ge -2 \wedge y - x \le 3 : & 0.1 \\ b \wedge 2x + z > 10 \wedge y - x \ge -2 \wedge y - x > 3 : & 0 \\ b \wedge 2x + z > 10 \wedge y - x < -2 : & 0 \\ \neg b \wedge 2x + z > 10 \wedge y - x \ge -2 \wedge y - x \le 3 : & 0.1 \\ \neg b \wedge 2x + z > 10 \wedge y - x \ge -2 \wedge y - x > 3 : & 0 \\ \neg b \wedge 2x + z > 10 \wedge y - x < -2 : & 0 \\ \neg b \wedge 2x + z > 10 \wedge y < -6 : & 0 \\ \neg b \wedge 2x + z > 10 \wedge y \ge -6 \wedge y > 4 : & 0 \\ \neg b \wedge 2x + z > 10 \wedge y \ge -6 \wedge y \le 4 : & y^2 \end{cases} \longrightarrow$$



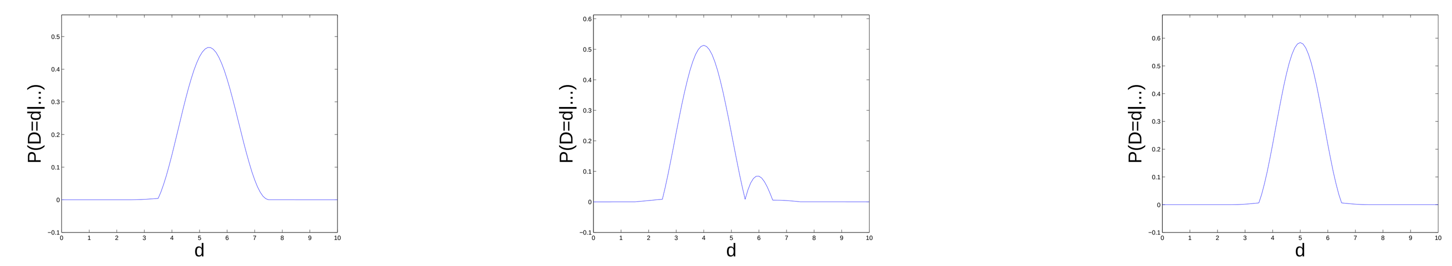
Example Applications

Robotics: Find exact posterior distance given complex sensor noise model:



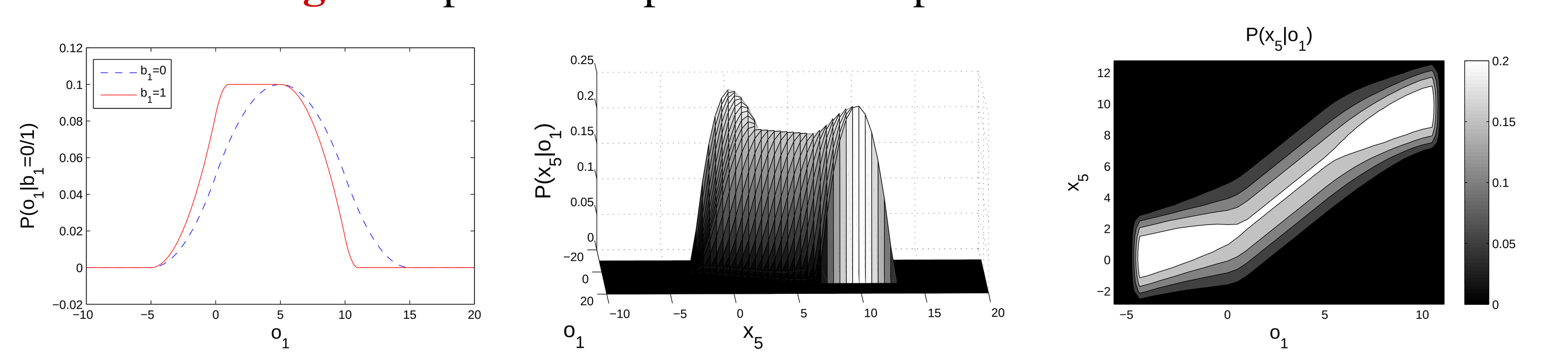


$\mathbb{E}[D|x_1 = 6, x_2 = 5] = 4.98$ $\mathbb{E}[D|x_1 = 8, x_2 = 5] = 6.0$ $\mathbb{E}[D|x_1 = 5, x_2 = 3, x_3 = 8] = 4.39$



$\mathbb{E}[D|x_1 = 5, x_2 = 5, x_3 = 6] = 4.98$ $\mathbb{E}[D|x_2 = 1, x_2 = 3, x_3 = 4, x_4 = 8] = 5.45$ $\mathbb{E}[D|x_1 = 5, x_2 = 4, x_3 = 6, x_4 = 5] = 4.89$

Radar Tracking: Compute exact posterior over position in skewed noise model:



Conclusion

First closed-form exact inference for complex continuous graphical models!