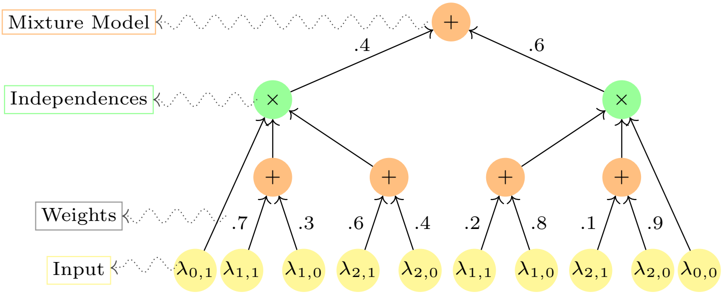
Robust Analysis of MAP Inference in Selective Sum-Product Networks

Julissa Villanueva Llerena Denis Deratani Mauá

Institute of Mathematics and Statistics
University of São Paulo
São Paulo, Brazil

Sum-Product Networks

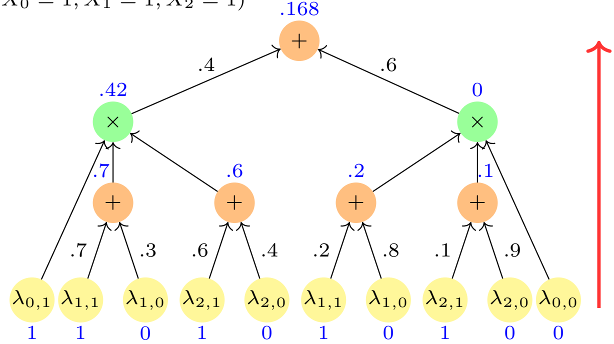
Neural Networks with clear Probabilistic Semantics
(non-linearity due to products)



Sum-Product Networks

- Tractable marginal probability inference
- Intractable Maximum-A-Posteriori (MAP) Inference

$\mathbb{P}(X_0 = 1, X_1 = 1, X_2 = 1)$

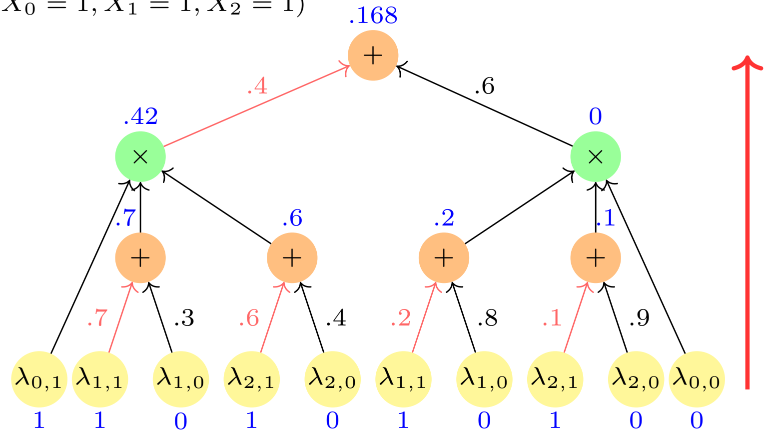


Sum-Product Networks Applications

Require (approximate) MAP Inference

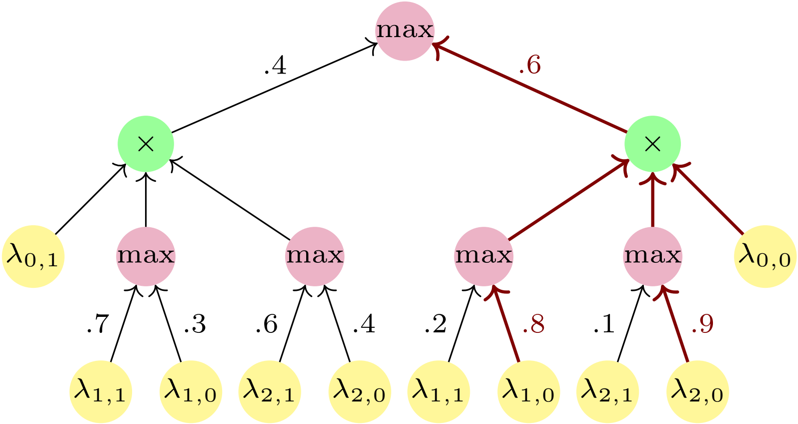
Selective Sum-Product Networks

$\mathbb{P}(X_0 = 1, X_1 = 1, X_2 = 1)$



MAP inference Selective Sum-Product Networks

$\max_{\mathbf{x}} \mathbb{P}(\mathbf{x}) : X_0 = 0, X_1 = 0, X_2 = 0$



Robust Analysis in Sum-Product Networks

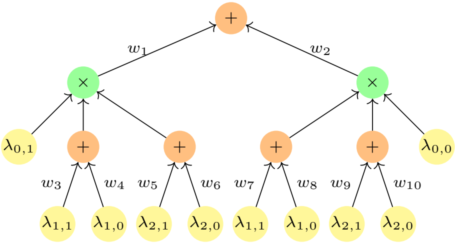
Goal: Distinguish between robust and non-robust MAP inferences

Robustness is assessed via imprecise mixtures:

- ϵ-contamination:
  $\mathcal{C}_{\mathbf{w}_i,\epsilon} = \left\{(1-\epsilon)\mathbf{w}_i + \epsilon\mathbf{v} : v_j \geq 0, \sum_j v_j = 1\right\}$
- Imprecise Dirichlet Model:
  $\mathcal{C}_{N,s} = \left\{\mathbf{w}_i : w_{ij} = \frac{N_j + s \cdot v_i}{N_i + s}, v_j \geq 0, \sum_j v_j = 1\right\}$

Robust Analysis in Sum-Product Netowrks

Credal Sum-Product Networks



$$
\begin{array}{ll}
0.36 \leq w_1 \leq 0.46 & w_1 + w_2 = 1 \\
0.63 \leq w_3 \leq 0.73 & w_3 + w_4 = 1 \\
0.54 \leq w_5 \leq 0.64 & w_5 + w_6 = 1 \\
0.18 \leq w_7 \leq 0.28 & w_7 + w_8 = 1 \\
0.09 \leq w_9 \leq 0.19 & w_9 + w_{10} = 1
\end{array}
$$

Figura: Credal SPN obtained by 0.1-contamination

Robust Analysis of MAP Inference in Selective Sum-Product Networks

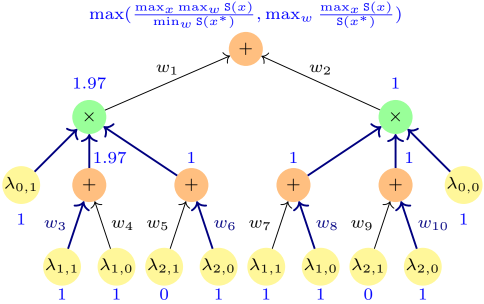
MAP configuration $\mathbf{x}^*$ is robust if:

$$\max_{\mathbf{x} \neq \mathbf{x}^*} \max_{\mathbf{w} \in \mathcal{C}} \left(\frac{\mathsf{S}_{\mathbf{w}}(\mathbf{x}, \mathbf{e})}{\mathsf{S}_{\mathbf{w}}(\mathbf{x}^*, \mathbf{e})} \right) < 1 \, .$$

I.e. $\mathbf{x}^*$ is still MAP inference in any perturbed SPN

Robust Analysis of MAP Inference in Selective Sum-Product Networks

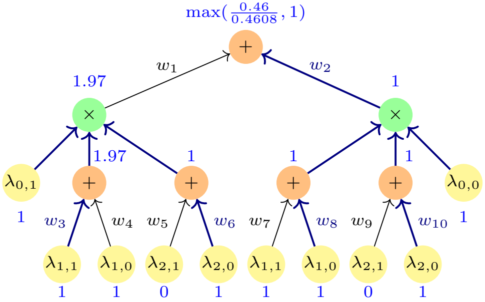
Let $\mathbf{x}^* : X_0 = 0, X_1 = 0$ be the MAP inference and $e : X_2 = 0$ be the evidence, we decide if $\mathbf{x}^*$ is robust for 0.1 contamination as follow:



$$0.36 \leq w_1 \leq 0.46 \qquad w_1 + w_2 = 1$$
$$0.63 \leq w_3 \leq 0.73 \qquad w_3 + w_4 = 1$$
$$0.54 \leq w_5 \leq 0.64 \qquad w_5 + w_6 = 1$$
$$0.18 \leq w_7 \leq 0.28 \qquad w_7 + w_8 = 1$$
$$0.09 \leq w_9 \leq 0.19 \qquad w_9 + w_{10} = 1$$

Robust Analysis of MAP Inference in Selective Sum-Product Networks

Let $\mathbf{x}^* : X_0 = 0, X_1 = 0$ be the MAP inference and $e : X_2 = 0$ be the evidence, we decide if $\mathbf{x}^*$ is robust for 0.1 contamination as follow:



Robust

$$0.36 \leq w_1 \leq 0.46 \qquad w_1 + w_2 = 1$$
$$0.63 \leq w_3 \leq 0.73 \qquad w_3 + w_4 = 1$$
$$0.54 \leq w_5 \leq 0.64 \qquad w_5 + w_6 = 1$$
$$0.18 \leq w_7 \leq 0.28 \qquad w_7 + w_8 = 1$$
$$0.09 \leq w_9 \leq 0.19 \qquad w_9 + w_{10} = 1$$

Experiments

Train	Learn the Selective SPNs from data using Peharz et al. 2014's algorithm
Valid	Choose s^*, ϵ^* and p^* that maximize the quality of robust predictions
Test	Use selected s^*, ϵ^* and p^* to compare performances of robust and non-robust instances on held-out data

Results: Multilabel Classification

Dataset		Accuracy			Exact Match		
		Robust	$\neg$*Robust*	ΔAcc	*Robust*	$\neg$*Robust*	ΔEM
Arts	$\epsilon*$	**0.88**	0.196	0.634	**0.833**	0.143	**0.69**
	$s*$	0.107	0.351	-0.244	0.089	0.247	-0.158
	$p*$	0.81	0.159	**0.651**	0.75	0.107	0.643
Business	$\epsilon*$	0.751	0.582	0.169	0.617	0.598	0.019
	$s*$	**0.781**	0.641	0.14	**0.642**	0.469	0.173
	$p*$	0.762	0.581	**0.181**	0.62	0.392	**0.228**
Emotions	$\epsilon*$	0.595	0.41	0.185	0.238	0.163	0.075
	$s*$	**0.686**	0.413	**0.273**	**0.308**	0.16	**0.148**
	$p*$	0.574	0.391	0.183	0.176	0.176	0
Flags	$\epsilon*$	0.917	0.468	0.449	0.5	0.118	0.382
	$s*$	0.917	0.468	0.449	0.5	0.1	**0.4**
	$p*$	0.917	0.468	0.449	0.5	0.118	0.382
Health	$\epsilon*$	**0.667**	0.557	0.11	0.5	0.416	0.084
	$s*$	0.637	0.482	**0.155**	0.537	0.304	**0.233**
	$p*$	0.655	0.552	0.103	**0.552**	0.409	0.143
Human	$\epsilon*$	0.203	-	-	0.146	-	-
	$s*$	0.203	-	-	0.146	-	-
	$p*$	**0.211**	0.198	0.013	**0.155**	0.14	0.015
Plant	$\epsilon*$	0.331	0.217	0.114	0.324	0.213	0.111
	$s*$	**0.367**	0.162	**0.205**	**0.362**	0.157	**0.205**
	$p*$	0.345	0.212	0.132	0.338	0.208	0.13
Scene	$\epsilon*$	0.857	0.277	0.58	0.857	0.212	0.645
	$s*$	**0.929**	0.293	0.636	**0.929**	0.293	0.636
	$p*$	0.923	0.276	**0.647**	0.923	0.211	**0.712**
Yeast	$\epsilon*$	**0.436**	0.419	0.017	0.071	0.098	-0.027
	$s*$	**0.436**	0.419	0.017	0.071	0.098	-0.027
	$p*$	0.425	0.203	**0.222**	**0.099**	0	**0.099**

Results: Density Estimation

Dataset		Exact Match			Hamming Score		
		Robust	$\neg$*Robust*	ΔEM	*Robust*	$\neg$*Robust*	ΔHS
Jester	$\epsilon*$	**0.259**	0.001	**0.258**	**0.913**	0.692	**0.221**
	$s*$	0.076	.001	0.075	0.755	0.692	0.063
	$p*$	0.016	0.001	0.015	0.858	0.687	0.171
NLTCS	$\epsilon*$	0.736	0.248	0.488	**0.934**	0.77	**0.164**
	$s*$	0.736	0.248	0.488	**0.934**	0.77	**0.164**
	$p*$	0.736	0.248	0.488	0.933	0.77	0.163
MSNBC	$\epsilon*$	**1**	0.247	**0.753**	**1**	0.783	**0.217**
	$s*$	0.464	0.215	0.249	0.875	0.769	0.106
	$p*$	0.727	0.247	0.48	0.92	0.783	0.137

See you at the Poster Session!