Global Sensitivity Analysis of MAP inference in Selective Sum-Product Networks

Julissa Villanueva Llerena, Denis Mauá

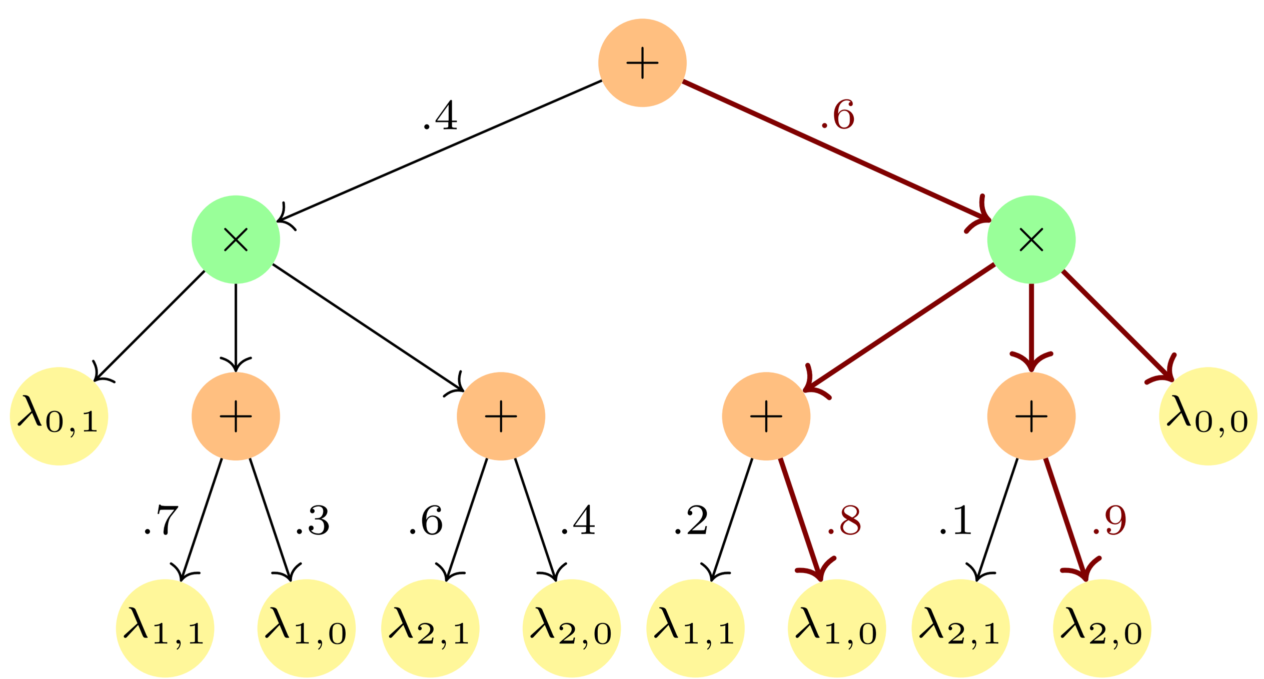
Global Sensitivity Analysis of MAP Inference in Selective Sum-Product Networks

Julissa Villanueva Llerena and Denis Deratani Mauá
Department of Computer Science, University of São Paulo, Brazil
jgville@ime.usp.br

1. Introduction

- Sum-Product Networks are deep generative probabilistic models with state-of-the-art performance in several machine learning tasks.
- Models learned from data can produce unreliable and overconfident inference on regions where data are scarce
- Qualitative Global Sensitivity Analysis consists detects whether prediction changes when model parameters are pertubed independently
- **Goal:** Tractable global sensitivity analysis of structured prediction

2. Selective Sum-Product Networks & MAP Inference



A **Selective SPN** is one of the following:

- An indicator variable $\lambda_{i,j}(X_i)$ assinging value 1 if $X_i = x_{ij}$ and 0 otherwise;
- A product $\prod_i S_i(x)$ of Selective SPNs S_i with disjoint scopes;
- A weighted sum $\sum_i w_i S_i(x)$ of selective SPNs S_i, where $\sum_i w_i = 1$, $w_i \geq 0$, and for each x at most one $S_i(x) > 0$.

Maximum a Posteriori (MAP) Inference

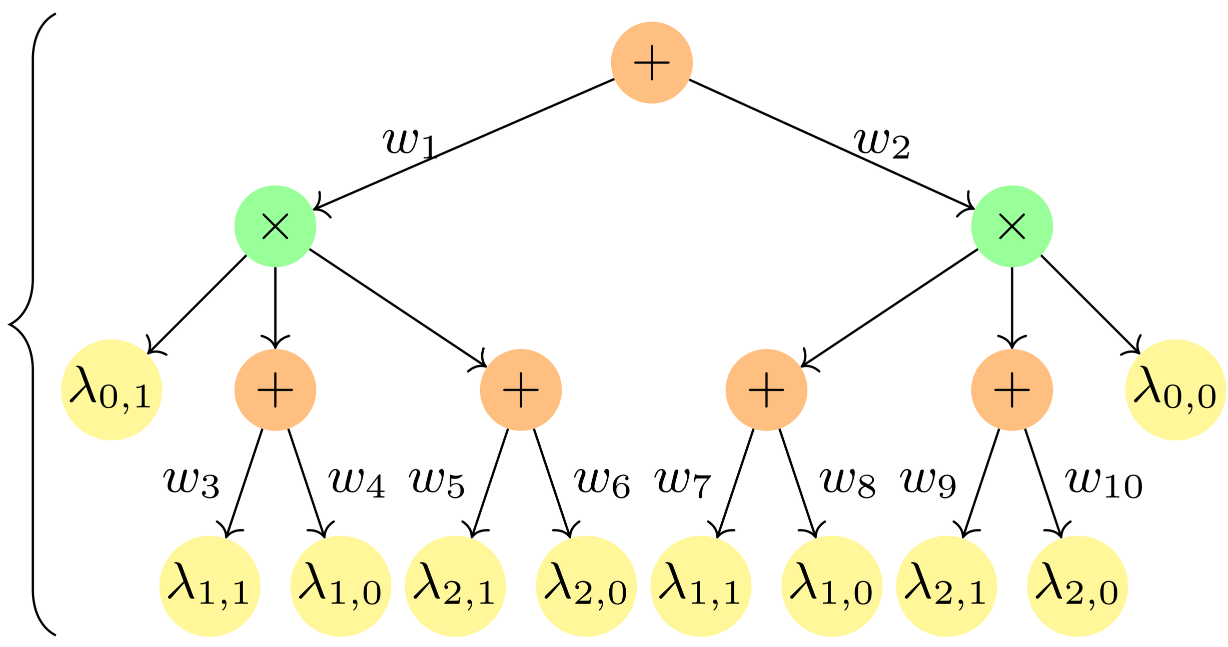
Given SPN representing joint distribution $\mathbb{P}(X, E)$, find $\mathbf{x}^* \in \text{argmax}_{\mathbf{x} \in val(\mathbf{X})} \mathbb{P}(\mathbf{x}|\mathbf{e})$
This is used for structured prediction. For Selective SPNs:

replace + with *max*

3. Global Sensivity Analysis

Let $\mathsf{S}_{\mathbf{w}}$ be an SPN whose weights are $\mathbf{w}$, Credal SPN is defined as a set of SPNs obtained by perturbing the weights $\mathbf{w}$ inside some fixed space $\mathcal{C}$, $\{\mathsf{S}_{\mathbf{w}} : \mathbf{w} \in \mathcal{C}\}$ obtained by:

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



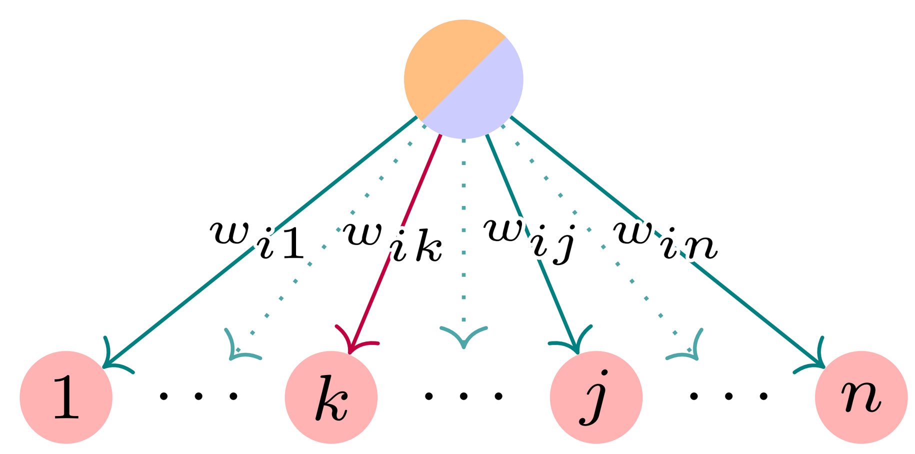
$0.36 \leq w_1 \leq 0.46 \qquad w_1 + w_2 = 1$
$0.63 \leq w_3 \leq 0.73 \qquad w_3 + w_4 = 1$
$0.54 \leq w_5 \leq 0.64 \qquad w_5 + w_6 = 1$
$0.18 \leq w_7 \leq 0.28 \qquad w_7 + w_8 = 1$
$0.09 \leq w_9 \leq 0.19 \qquad w_9 + w_{10} = 1$

4. Robust MAP Inference

A MAP inference $\mathbf{x}^*$ is robust with respect to a credal SPN $\{\mathsf{S}_{\mathbf{w}} : \mathbf{w} \in \mathcal{C}\}$ if:

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

$$V^i = \begin{cases} i \text{ is consistent leaf,} & 1 \\ i \text{ is inconsistent leaf,} & 0 \\ i \text{ is product node,} & \prod_j V^j \\ i \text{ is sum node ,} & \max\left\{\max_{j \in ch(i), j \neq k}\left(\max_{\mathbf{w}_i \in \mathcal{C}_i} \frac{w_{ij} \max_{\mathbf{x}} \mathsf{S}^j_{\mathbf{w}_j}(\mathbf{x},\mathbf{e})}{w_{ik} \mathsf{S}^k_{\mathbf{w}_k}(\mathbf{x}^*,\mathbf{e})}\right), V^k\right\} \end{cases}$$



5. Results

Dataset		Accuracy					Exact Match				
		Robust	*%I*	*¬Robust*	*%I*	ΔAcc	*Robust*	*%I*	*¬Robust*	*%I*	ΔEM
Arts	$\epsilon*$	**0.88**	1.2	0.196	98.8	0.634	**0.833**	1.2	0.143	98.8	**0.69**
	$s*$	0.107	51.94	0.351	48.06	-0.244	0.089	25.8	0.247	74.2	-0.158
	$p*$	0.81	6.95	0.159	93.05	**0.651**	0.75	6.95	0.107	93.05	0.643
Business	$\epsilon*$	0.751	72.73	0.582	27.27	0.169	0.617	65.24	0.598	34.76	0.019
	$s*$	**0.781**	45.32	0.641	54.68	0.14	**0.642**	45.32	0.469	54.68	0.173
	$p*$	0.762	68.4	0.581	31.6	**0.181**	0.62	68.4	0.392	31.6	**0.228**
Emotions	$\epsilon*$	0.595	17.65	0.41	82.35	0.185	0.238	17.65	0.163	82.35	0.075
	$s*$	**0.686**	10.92	0.413	89.08	**0.273**	**0.308**	10.92	0.16	89.08	**0.148**
	$p*$	0.574	28.57	0.391	71.43	0.183	0.176	28.57	0.176	71.43	0
Flags	$\epsilon*$	0.917	10.53	0.468	89.47	0.449	0.5	57.89	0.118	42.11	0.382
	$s*$	0.917	10.53	0.468	89.47	0.449	0.5	47.37	0.1	52.63	**0.4**
	$p*$	0.917	10.53	0.468	89.47	0.449	0.5	10.53	0.118	89.47	0.382

Dataset		Accuracy					Exact Match				
		Robust	*%I*	*¬Robust*	*%I*	ΔAcc	*Robust*	*%I*	*¬Robust*	*%I*	ΔEM
Health	$\epsilon*$	**0.667**	0.11	0.557	99.89	0.11	0.5	0.11	0.416	99.89	0.084
	$s*$	0.637	47.88	0.482	52.12	**0.155**	0.537	47.88	0.304	52.12	**0.233**
	$p*$	0.655	4.72	0.552	95.28	0.103	**0.552**	4.72	0.409	95.28	0.143
Human	$\epsilon*$	0.203	100	-	0	-	0.146	100	-	0	-
	$s*$	0.203	100	-	0	-	0.146	100	-	0	-
	$p*$	**0.211**	42.6	0.198	57.4	0.013	**0.155**	42.6	0.14	57.4	0.015
Plant	$\epsilon*$	0.331	37.76	0.217	62.24	0.114	0.324	37.76	0.213	62.24	0.111
	$s*$	**0.367**	47.96	0.162	52.04	**0.205**	**0.362**	47.96	0.157	52.04	**0.205**
	$p*$	0.345	36.22	0.212	63.76	0.132	0.338	36.22	0.208	63.76	0.13
Scene	$\epsilon*$	0.857	5.83	0.277	94.17	0.58	0.857	5.83	0.212	94.17	0.645
	$s*$	**0.929**	2.92	0.293	0.636	97.08	**0.929**	2.92	0.293	97.08	0.636
	$p*$	0.923	5.42	0.276	94.58	**0.647**	0.923	5.42	0.211	94.58	**0.712**

6. References and Acknowledgements

[1] Julissa Villanueva and Denis Mauá. Robust analysis of map inference in selective sum-product networks. In *International Symposium on Imprecise Probabilities: Theories and Applications*, pages 430–440, 2019.

[2] Robert Peharz, Robert Gens, and Pedro Domingos. Learning selective sum-product networks. In *Workshop LTPM*, 2014.

[3] D. Mauá, D. Conaty, F. Cozman, K. Poppenhaeger, and C. de Campos. Robustifying sum-product networks. *International Journal of Approximate Reasoning*, 2018.

[4] J. De Bock, C. De Campos, and A. Antonucci. Global sensitivity analysis for map inference in graphical models. In *Proceedings of NIPS*, 2014.

CAPES