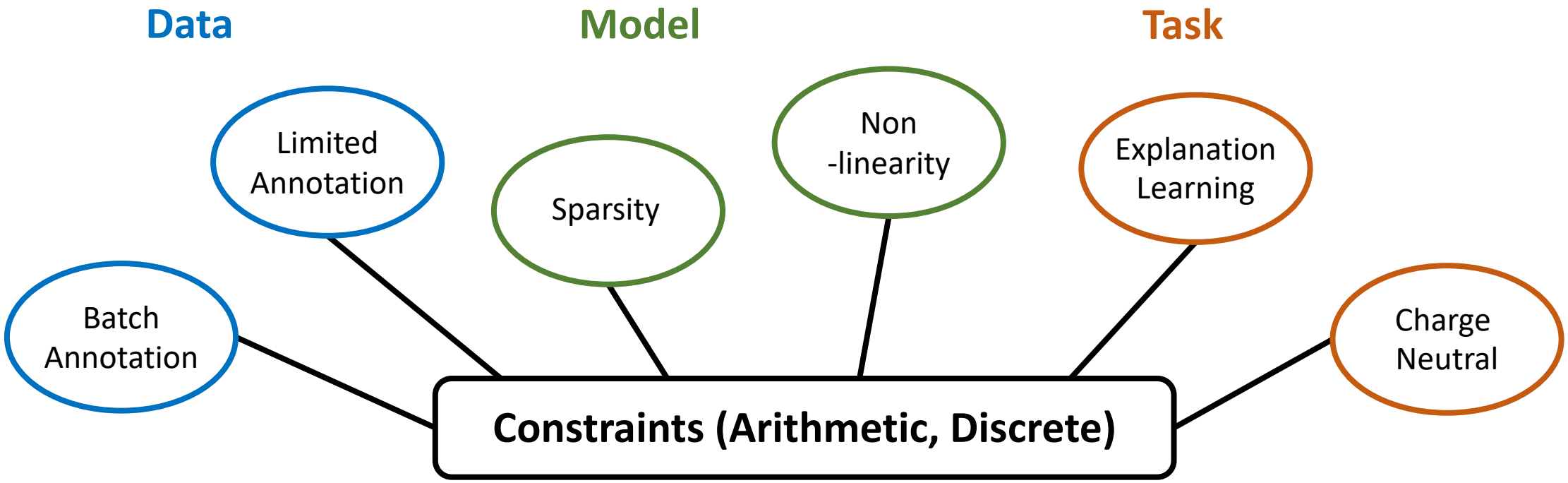
Learning and Probabilistic Inference with Constraints and its Applications

Zhe Zeng

Advisor: Prof. Guy Van den Broeck

Overview

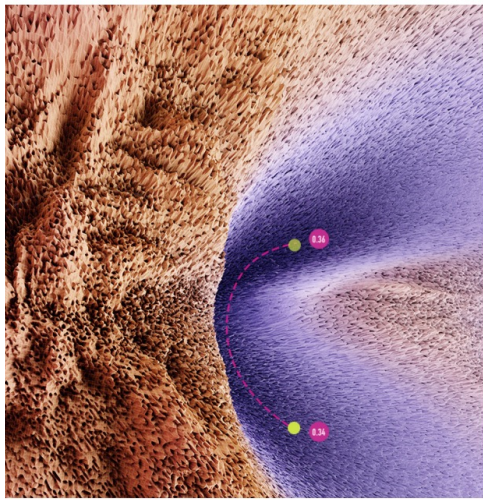


- **Goal**: To enable machine learning models to perform learning and probabilistic inference under constraints

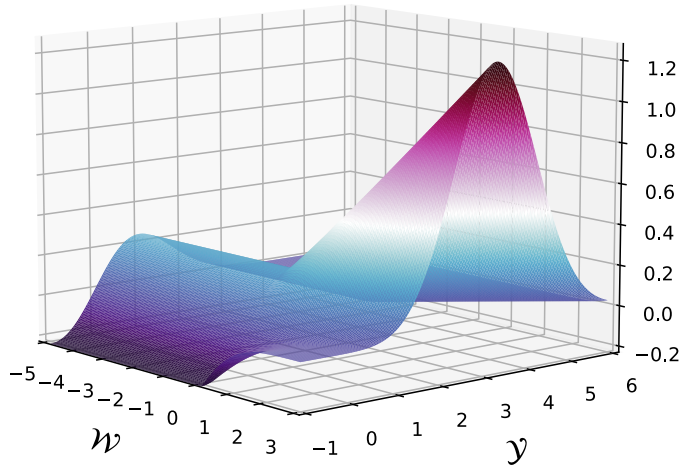
Bayesian Deep Learning [arithmetic constraints]

- **Goal**: *Marginalization* over the weight space
  - to aggregate models with low loss

Loss Surface



Predictions and Uncertainty



$$p(y \mid \boldsymbol{x}) = \int p(y \mid \boldsymbol{x}, \boldsymbol{w})\, p(\boldsymbol{w} \mid \mathcal{D})\, d\boldsymbol{w}$$

$$\mathbb{E}_{p(y|\boldsymbol{x})}[y] = \int y\, p(y \mid \boldsymbol{x})\, dy$$

under constraints from ReLU:

$$\boldsymbol{x} \cdot \boldsymbol{w} \geq 0$$

- **Solution**: efficiently and effectively by *Weighted Model Integration*

T. Garipov, P. Izmailov, D. Podoprikhin, D. Vetrov, A.G. Wilson. *Loss Surfaces, Mode Connectivity, and Fast Ensembling of DNNs*, NeurIPS 2018

Z. Zeng, P. Morettin, F. Yan, A. Vergari, G. Van den Broeck. *Probabilistic inference with algebraic constraints: Theoretical limits and practical approximations*, NeurIPS 2020

Z. Zeng, G. Van den Broeck. *Collapsed Inference for Bayesian Deep Learning*, Beyond Bayes, ICML Workshop, 2022

Gradient Estimator for k-Subset [discrete constraints]

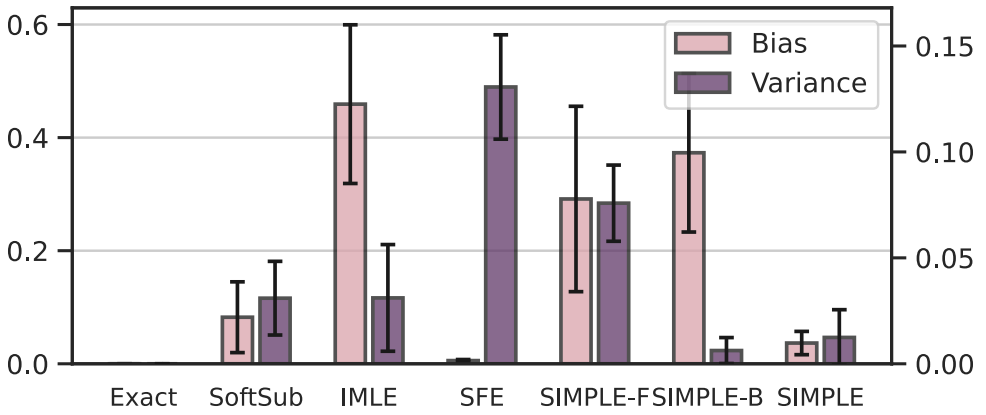
- **Goal**: Modeling a distribution of *k*-subsets of elements $p_{\boldsymbol{\theta}}(\mathbf{z} \mid \sum_i z_i = k)$

Example. Learn to Explain

Key Words (*k* = 10)	Taste Score
a lite bodied beer with a pleasant taste. was like a reddish color. a little like wood and caramel with a hop finish. has a sort of fruity flavor like grapes or cherry that is sort of buried in there. mouth feel was lite, sort of bubbly. not hard to down, though a bit harder then one would expect given the taste.	0.7

- **Solution**:

SIMPLE that computes exact samples and exact derivatives



Kareem Ahmed*, Zhe Zeng*, Mathias Niepert, Guy Van den Broeck. *SIMPLE: A Gradient Estimator for k-Subset Sampling*, ICLR 2023

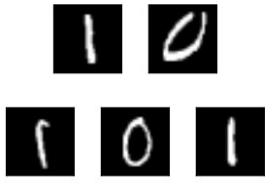
Weakly Supervised Learning [discrete constraints]

- **Goal**: To train classifiers under weak supervisions

classical

$\boldsymbol{x}$	y
	0
	0
	1
	1

learning from label proportions

$\{\boldsymbol{x}_i\}_{i=1}^k$	$\tilde{y} = \sum y_i / k$
	0
	1/3
	3/5

multiple instance learning

$\{\boldsymbol{x}_i\}_{i=1}^k$	$\tilde{y} = \max\{y_i\}$
	0
	1
	1

learning from positive and unlabeled data

$\boldsymbol{x}$	$\tilde{y}$
	?
	1
	?
	?

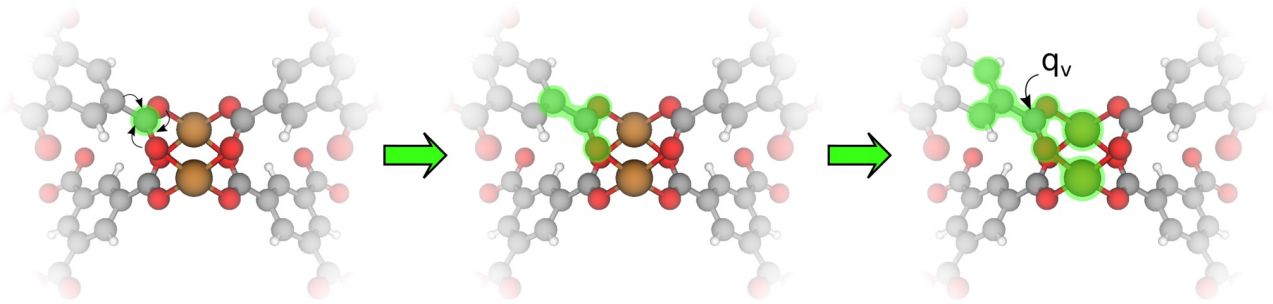
- **Solution**: To use *probability* of these *constraints on label counts* being satisfied as training objectives

Vinay Shukla, Zhe Zeng, Kareem Ahmed, Guy Van den Broeck. *A Unified Approach to Count-Based Weakly-Supervised Learning.*

Ongoing and Future Work

- To build machine learning models
  - that can deal with complicated constraints
  - and deliver accurate and efficient inference

- Applications to various fields
  - computational chemistry: metal-organic frameworks
  - more ……



- **Thank you for your time!**

- Q & A
- Email: zhezeng@cs.ucla.edu