1oML

Neural Learning of One-of-Many Solutions for Combinatorial Problems in Structured Output Spaces

Yatin Nandwani, Deepanshu Jindal, Mausam & Parag Singla
Department of Computer Science & Engineering
Indian Institute of Technology Delhi, INDIA

Neural Networks for Symbolic Reasoning

System 2 Deep Learning (Bengio, 2019) – need for modern DL

Neural networks with ability to reason over and above perception.

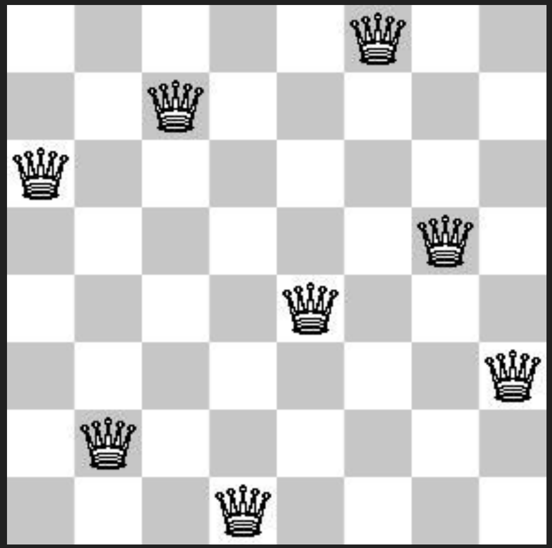
Neural Networks for Symbolic Reasoning

System 2 Deep Learning (Bengio, 2019) – need for modern DL

Neural networks with ability to reason over and above perception.

5	3			7				
6			1	9	5			
	9	8					6	
8				6				3
4			8		3			1
7				2				6
	6					2	8	
			4	1	9			5
				8			7	9

Sudoku



N-Queens

Neural Networks for Symbolic Reasoning

System 2 Deep Learning (Bengio, 2019) – need for modern DL

Neural networks with ability to reason over and above perception.

- NLM: Neural Logic Machines [Dong et al., 2019]
-

Neural Networks for Symbolic Reasoning

System 2 Deep Learning (Bengio, 2019) – need for modern DL

Neural networks with ability to reason over and above perception.

- NLM: Neural Logic Machines [Dong et al., 2019]
- RRN: Recurrent Relational Networks [Palm et al., 2018]
-

Neural Networks for Symbolic Reasoning

System 2 Deep Learning (Bengio, 2019) – need for modern DL

Neural networks with ability to reason over and above perception.

- NLM: Neural Logic Machines [Dong et al., 2019]
- RRN: Recurrent Relational Networks [Palm et al., 2018]
- SATNet: Bridging deep learning and logical reasoning … [Wang et al., 2019]

Neural Networks for Symbolic Reasoning

System 2 Deep Learning (Bengio, 2019) – need for modern DL

Neural networks with ability to reason over and above perception.

- NLM: Neural Logic Machines [Dong et al., 2019]
- RRN: Recurrent Relational Networks [Palm et al., 2018]
- SATNet: Bridging deep learning and logical reasoning … [Wang et al., 2019]

Discover and implicitly encode logical relationships in structured output spaces

Solution Multiplicity

Solution Multiplicity

2	9	5	7	4	3	8	6	1
4	3	1	8	6	5	9		
8	7	6	1	9	2	5	4	3
3	8	7	4	5	9	2	1	6
6	1	2	3	8	7	4	9	5
5	4	9	2	1	6	7	3	8
7	6	3	5	3	4	1	8	9
9	2	8	6	7	1	3	5	4
1	5	4	9	3	8	6		

Solution Multiplicity

2	9	5	7	4	3	8	6	1
4	3	1	8	6	5	9		
8	7	6	1	9	2	5	4	3
3	8	7	4	5	9	2	1	6
6	1	2	3	8	7	4	9	5
5	4	9	2	1	6	7	3	8
7	6	3	5	3	4	1	8	9
9	2	8	6	7	1	3	5	4
1	5	4	9	3	8	6		

2	9	5	7	4	3	8	6	1
4	3	1	8	6	5	9	2	7
8	7	6	1	9	2	5	4	3
3	8	7	4	5	9	2	1	6
6	1	2	3	8	7	4	9	5
5	4	9	2	1	6	7	3	8
7	6	3	5	3	4	1	8	9
9	2	8	6	7	1	3	5	4
1	5	4	9	3	8	6	7	2

Solution Multiplicity

Many correct solutions for any given input.

2	9	5	7	4	3	8	6	1
4	3	1	8	6	5	9		
8	7	6	1	9	2	5	4	3
3	8	7	4	5	9	2	1	6
6	1	2	3	8	7	4	9	5
5	4	9	2	1	6	7	3	8
7	6	3	5	3	4	1	8	9
9	2	8	6	7	1	3	5	4
1	5	4	9	3	8	6		

2	9	5	7	4	3	8	6	1
4	3	1	8	6	5	9	2	7
8	7	6	1	9	2	5	4	3
3	8	7	4	5	9	2	1	6
6	1	2	3	8	7	4	9	5
5	4	9	2	1	6	7	3	8
7	6	3	5	3	4	1	8	9
9	2	8	6	7	1	3	5	4
1	5	4	9	3	8	6	7	2

2	9	5	7	4	3	8	6	1
4	3	1	8	6	5	9	7	2
8	7	6	1	9	2	5	4	3
3	8	7	4	5	9	2	1	6
6	1	2	3	8	7	4	9	5
5	4	9	2	1	6	7	3	8
7	6	3	5	3	4	1	8	9
9	2	8	6	7	1	3	5	4
1	5	4	9	3	8	6	2	7

Solution Multiplicity

Many correct solutions for any given input.

Interested in any one solution with no preference.

2	9	5	7	4	3	8	6	1
4	3	1	8	6	5	9		
8	7	6	1	9	2	5	4	3
3	8	7	4	5	9	2	1	6
6	1	2	3	8	7	4	9	5
5	4	9	2	1	6	7	3	8
7	6	3	5	3	4	1	8	9
9	2	8	6	7	1	3	5	4
1	5	4	9	3	8	6		

2	9	5	7	4	3	8	6	1
4	3	1	8	6	5	9	2	7
8	7	6	1	9	2	5	4	3
3	8	7	4	5	9	2	1	6
6	1	2	3	8	7	4	9	5
5	4	9	2	1	6	7	3	8
7	6	3	5	3	4	1	8	9
9	2	8	6	7	1	3	5	4
1	5	4	9	3	8	6	7	2

2	9	5	7	4	3	8	6	1
4	3	1	8	6	5	9	7	2
8	7	6	1	9	2	5	4	3
3	8	7	4	5	9	2	1	6
6	1	2	3	8	7	4	9	5
5	4	9	2	1	6	7	3	8
7	6	3	5	3	4	1	8	9
9	2	8	6	7	1	3	5	4
1	5	4	9	3	8	6	2	7

Solution Multiplicity

- Most neural models - trained and tested over unique solution puzzles.
- Completely ignore the issue of solution multiplicity.

Solution Multiplicity

- Most neural models - trained and tested over unique solution puzzles.
- Completely ignore the issue of solution multiplicity.

Recurrent Relation Network – SOTA neural Sudoku solver

- Solves 96% unique solution puzzles
- But only 24% multiple solution puzzles

Solution Multiplicity

- Explicit modeling required to represent this solution multiplicity.
- Real world reasoning problems may have multiple solutions.

Overview

- Problem formulation – ***One*** of **M**any **L**earning (**1oML**)
-

Overview

- Problem formulation – ***One of Many Learning* (1oML)**
- Naïve Solutions: *NaïveLoss*, *CC-Loss*
-

Overview

- Problem formulation – ***One* of Many Learning (1oML)**
- Naïve Solutions: *NaïveLoss*, *CC-Loss*
- Multiplicity Aware Loss Function: $\boldsymbol{L_w}$
-

Overview

- Problem formulation – ***One of Many Learning*** (**1oML**)
- Naïve Solutions: *NaïveLoss*, *CC-Loss*
- Multiplicity Aware Loss Function: $\boldsymbol{L_w}$
- Optimization Techniques
  - **Greedy** – *MinLoss*
  - **Exploration Based**: *IExplR* and *SelectR*
-

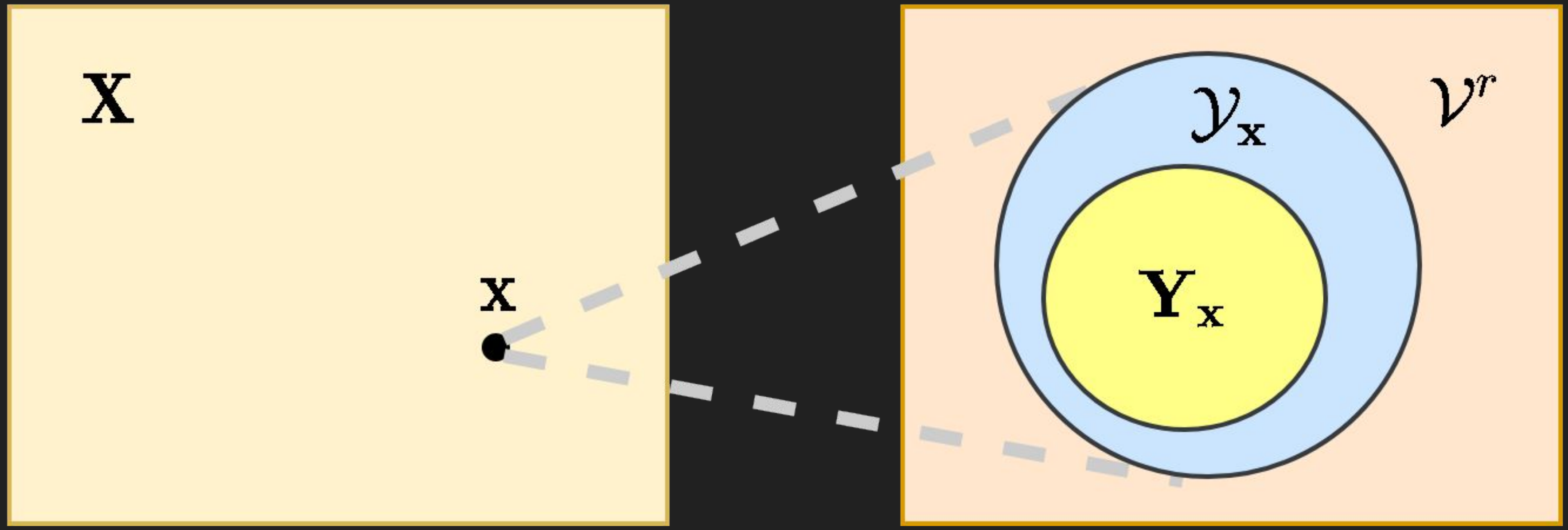
Overview

- Problem formulation – ***One* of Many Learning (1oML)**
- Naïve Solutions: *NaïveLoss*, *CC-Loss*
- Multiplicity Aware Loss Function: $\boldsymbol{L_w}$
- Optimization Techniques
    - **Greedy** – *MinLoss*
    - **Exploration Based**: *IExplR* and *SelectR*
- Experiments
    - Three domains
    - Two reasoning models

One of Many Learning (1oML)

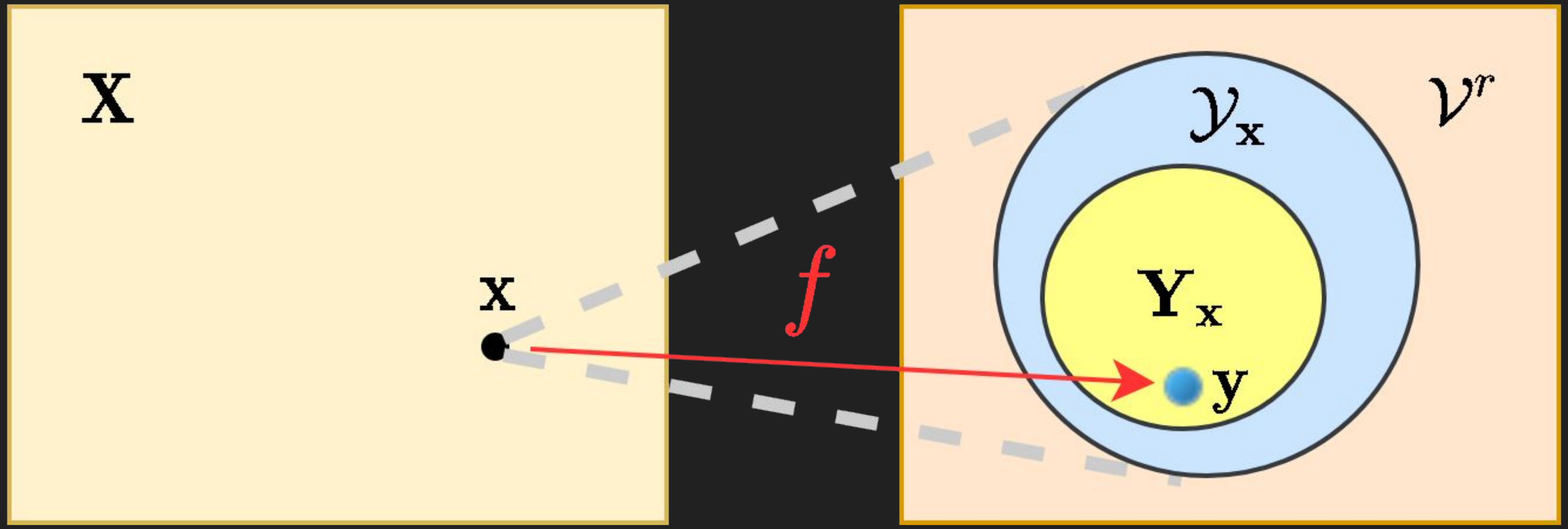
$$\mathbb{D} = \{(\mathbf{x}_i, \mathbf{Y}_{\mathbf{x}_i})\}_{i=1}^{m}$$

One of Many Learning (1oML)



$$\mathbb{D} = \{(\mathbf{x}_i, \mathbf{Y}_{\mathbf{x}_i})\}_{i=1}^{m}$$

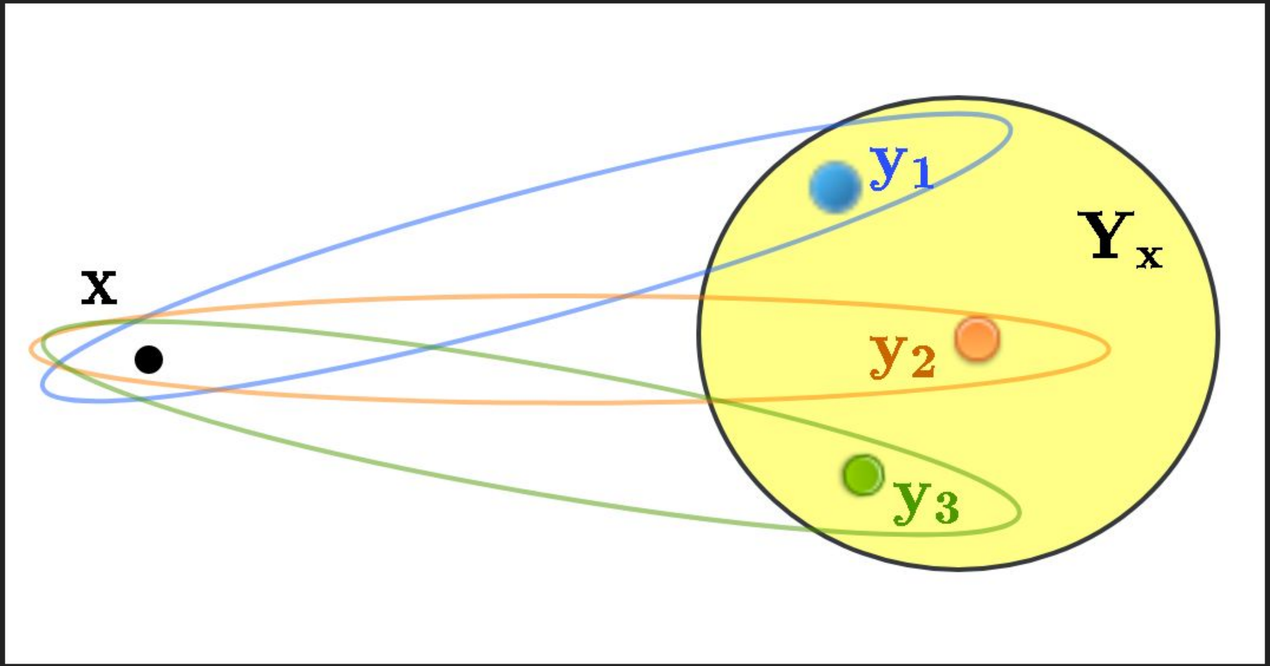
One of Many Learning (1oML)



$$\mathbb{D} = \{(\mathbf{x}_i, \mathbf{Y}_{\mathbf{x}_i})\}_{i=1}^{m}$$

Naïve Loss

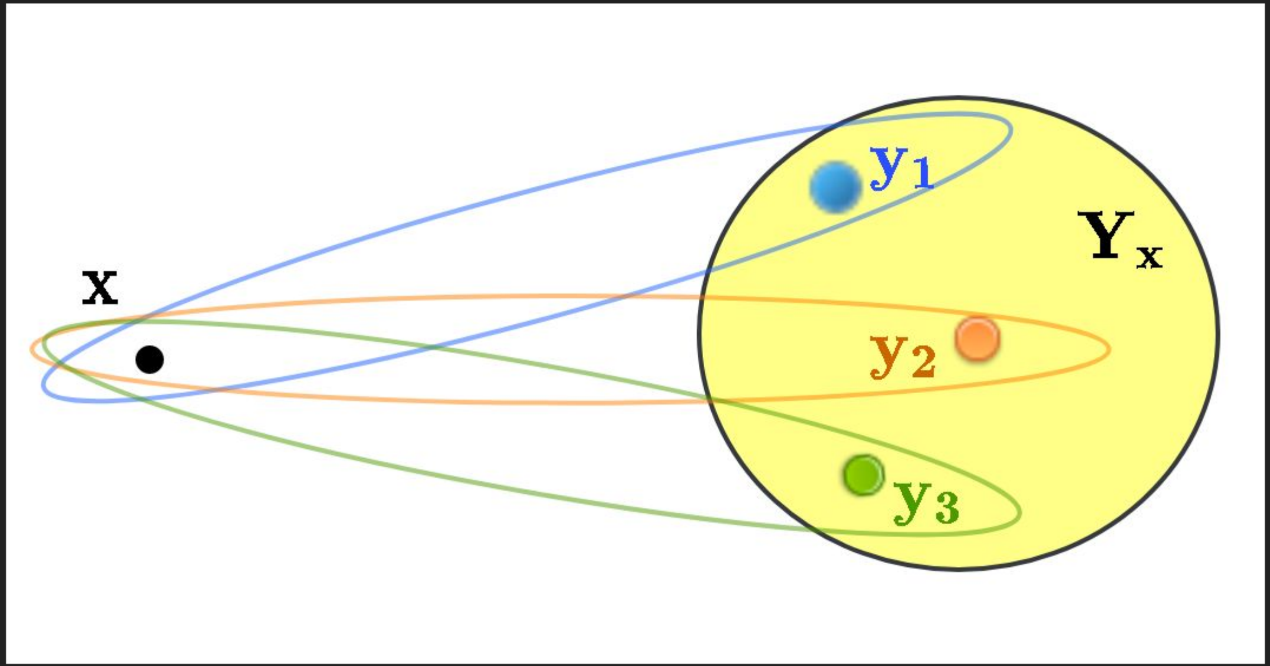
- Parameterize f as M_Θ



Naïve Loss

- Parameterize f as M_Θ

$$L(\Theta) = \sum_{\mathrm{i}=1}^{m} \sum_{\mathbf{y}_{\mathrm{ij}} \in \mathbf{Y}_{\mathbf{x}_\mathrm{i}}} l_\Theta(\hat{\mathbf{y}}_\mathrm{i}, \mathbf{y}_\mathrm{ij})$$



- Penalizes the model even if prediction is correct!

Reformulating Loss

Reformulating Loss

$$L_{\mathbf{w}}(\Theta, \mathbf{w}) = \sum_{\mathbf{i}=1}^{m} \sum_{\mathbf{y_{ij}} \in \mathbf{Y_{x_i}}} w_{\mathbf{ij}} l_{\Theta}(\hat{\mathbf{y}}_{\mathbf{i}}, \mathbf{y_{ij}})$$

$$s.t.\ w_{\mathbf{ij}} \in \{0, 1\}\ \forall \mathbf{i}, \forall \mathbf{j}\ and\ \sum_{\mathbf{j}=1}^{|\mathbf{Y_{x_i}}|} w_{\mathbf{ij}} = 1, \forall \mathbf{i} = 1 \ldots m$$

Greedy Optimization: ***MinLoss***

Greedily chooses w for each example based on current Θ

Greedy Optimization: *MinLoss*

Greedily chooses w for each example based on current Θ

$$w_{\mathbf{ij}} = 1\left\{\mathbf{y}_{\mathbf{ij}} = \underset{\mathbf{y}\in\mathbf{Y}_{\mathbf{x}_\mathbf{i}}}{argmin}\; l_{\Theta}\left(\hat{\mathbf{y}}_{\mathbf{i}}, \mathbf{y}\right)\right\}, \forall \mathbf{i} = 1 \ldots m$$

Greedy Optimization: *MinLoss*

Greedily chooses w for each example based on current Θ

$$w_{\mathbf{ij}} = 1\left\{\mathbf{y_{ij}} = \underset{\mathbf{y}\in\mathbf{Y_{x_i}}}{argmin}\ l_{\Theta}\left(\hat{\mathbf{y}}_{\mathbf{i}}, \mathbf{y}\right)\right\}, \forall \mathbf{i} = 1 \ldots m$$

Locally optimal choices might not be globally optimal.

Exploration based Optimization

Select non-greedy targets with non-zero probability

IExplR:

SelectR:

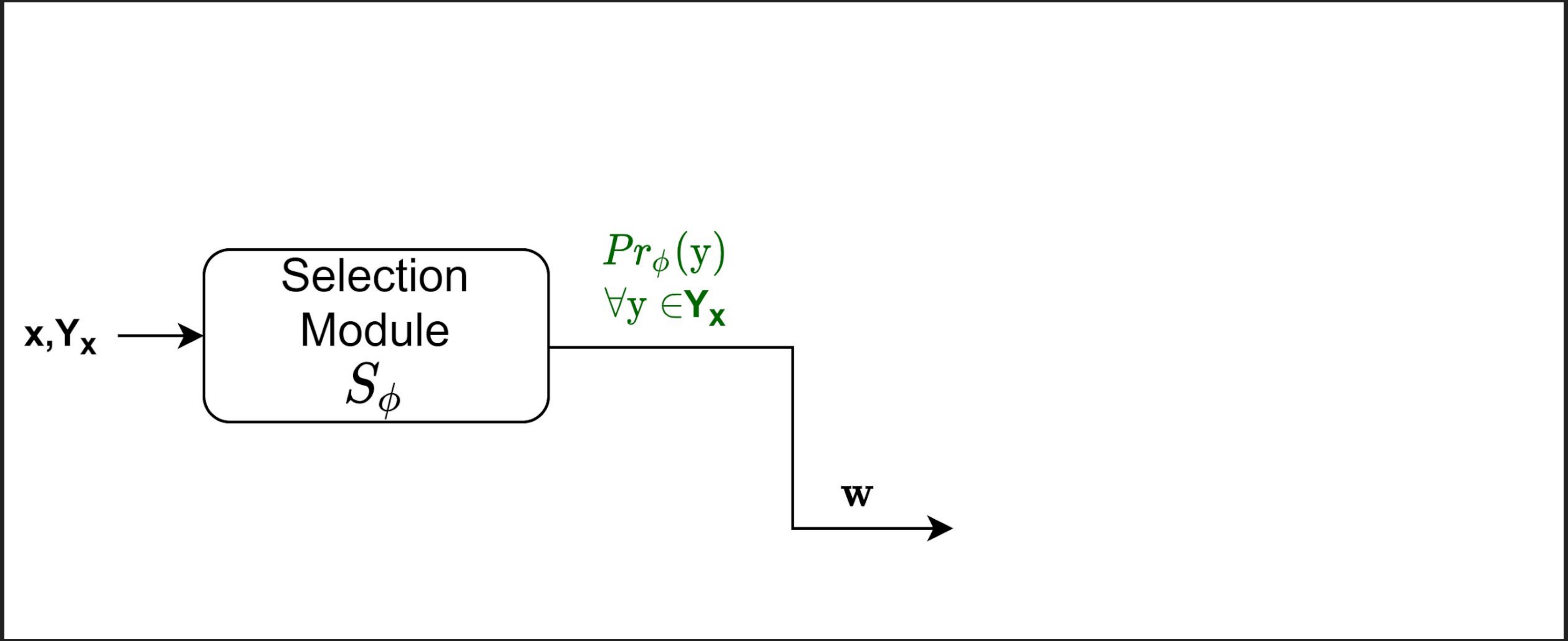
Exploration based Optimization

Select non-greedy targets with non-zero probability

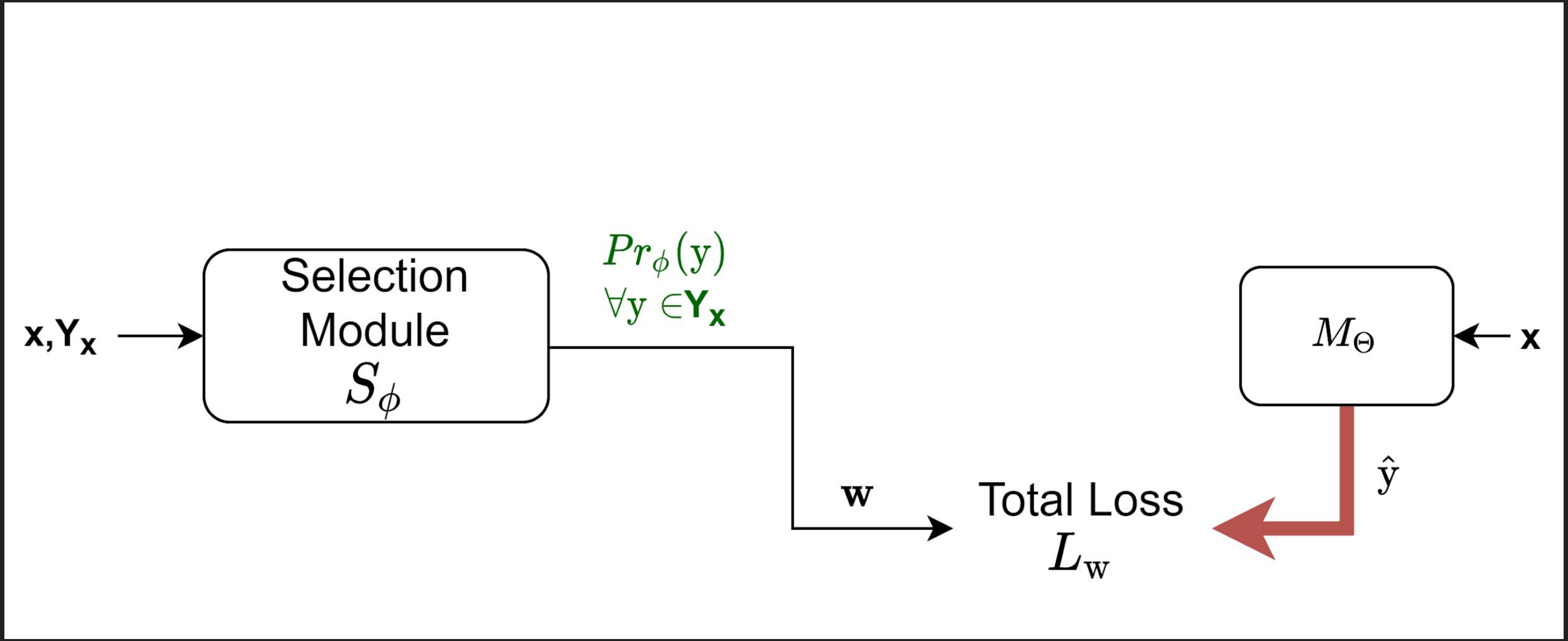
IExplR: Use $Pr(\mathbf{y_{ij}} | \mathbf{x}; \Theta)$ as exploration probability

SelectR: Use an RL agent to get exploration probability

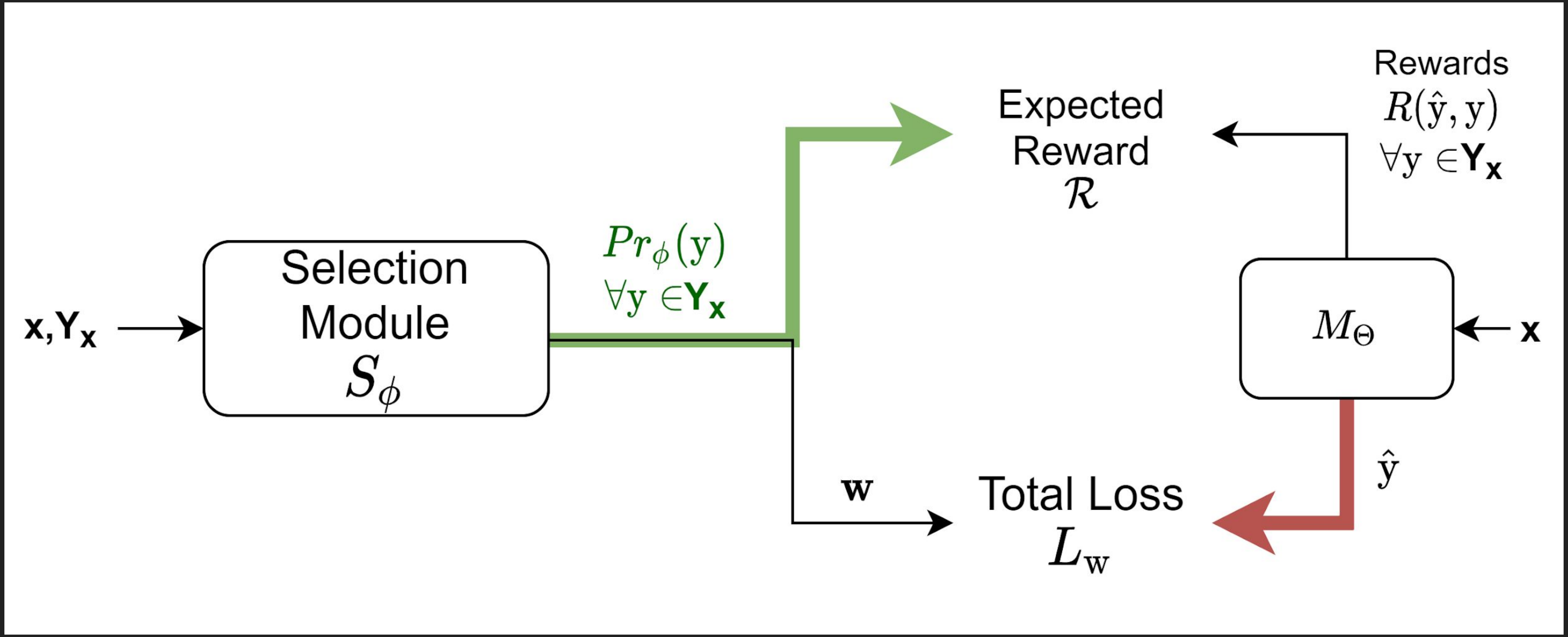
Exploration based Optimization: *SelectR*



Exploration based Optimization: *SelectR*



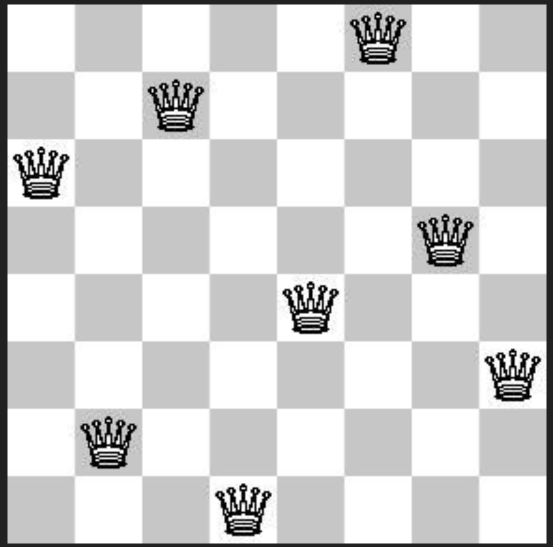
Exploration based Optimization: *SelectR*



Experiments

Tasks

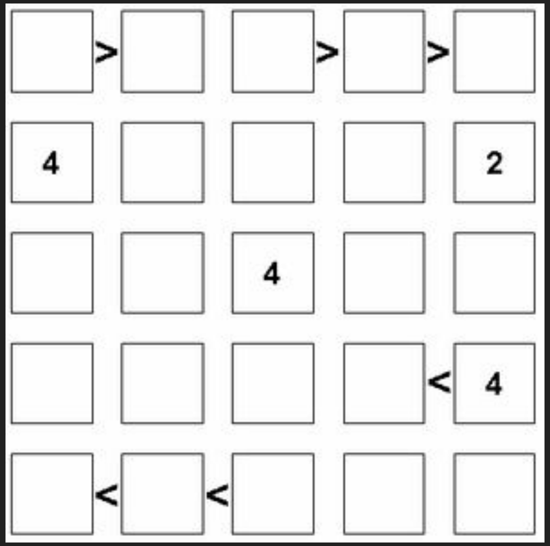
- NQueens



Experiments

Tasks

- NQueens
- Futoshiki



Experiments

Tasks

- NQueens
- Futoshiki
- Sudoku

5	3			7				
6			1	9	5			
	9	8					6	
8				6				3
4			8		3			1
7				2				6
	6					2	8	
			4	1	9			5
				8			7	9

Experiments

Tasks

- NQueens
- Futoshiki
- Sudoku

NLM

5	3			7				
6			1	9	5			
	9	8					6	
8				6				3
4			8		3			1
7				2				6
	6					2	8	
			4	1	9			5
				8			7	9

Experiments

Tasks

- NQueens
- Futoshiki
- Sudoku

NLM

RRN

5	3			7				
6			1	9	5			
	9	8					6	
8				6				3
4			8		3			1
7				2				6
	6					2	8	
			4	1	9			5
				8			7	9

Results

	Naïve	Unique	Random	CC-Loss	MinLoss	IExplR	SelectR
NQueens							
Futoshiki							
Sudoku							

Results

		Naïve	Unique	Random	CC-Loss	MinLoss	IExplR	SelectR
NQueens	Overall	68.04	73.72	70.94				
Futoshiki	Overall	52.96	55.53	52.70				
Sudoku	Overall	48.49	77.79	50.59				

Results

		Naïve	Unique	Random	CC-Loss	MinLoss	IExplR	SelectR
NQueens	**Overall**	68.04	73.72	70.94	75.39	77.28	77.70	**79.72**
Futoshiki	**Overall**	52.96	55.53	52.70	75.59	75.18	76.33	**76.40**
Sudoku	**Overall**	48.49	77.79	50.59	82.42	82.59	84.46	**85.21**

Results

		Naïve	Unique	Random	CC-Loss	MinLoss	IExplR	SelectR
NQueens	**OS**	70.59	75.09	72.91	75.31	77.29	77.35	**79.73**
	MS	55.34	66.85	61.13	75.76	77.22	79.46	**79.68**
	Overall	68.04	73.72	70.94	75.39	77.28	77.70	**79.72**
Futoshiki	**OS**							
	MS							
	Overall	52.96	55.53	52.70	75.59	75.18	76.33	**76.40**
Sudoku	**OS**							
	MS							
	Overall	48.49	77.79	50.59	82.42	82.59	84.46	**85.21**

Results

		Naïve	Unique	Random	CC-Loss	MinLoss	IExplR	SelectR
NQueens	OS	70.59	75.09	72.91	75.31	77.29	77.35	**79.73**
	MS	55.34	66.85	61.13	75.76	77.22	79.46	**79.68**
	Overall	68.04	73.72	70.94	75.39	77.28	77.70	**79.72**
Futoshiki	OS	65.59	67.63	65.49	77.68	76.78	**78.15**	78.01
	MS	14.99	19.13	14.22	69.30	70.35	70.88	**71.57**
	Overall	52.96	55.53	52.70	75.59	75.18	76.33	**76.40**
Sudoku	OS	87.85	**89.19**	87.53	88.26	88.25	88.73	88.69
	MS	9.13	66.39	13.65	76.58	76.93	80.19	**81.73**
	Overall	48.49	77.79	50.59	82.42	82.59	84.46	**85.21**

Questions?

Poster Session 11

6th May 2021

9:00 am to 11:00 am PDT

Resources:

https://sites.google.com/view/yatinnandwani/1oml

Thanks!