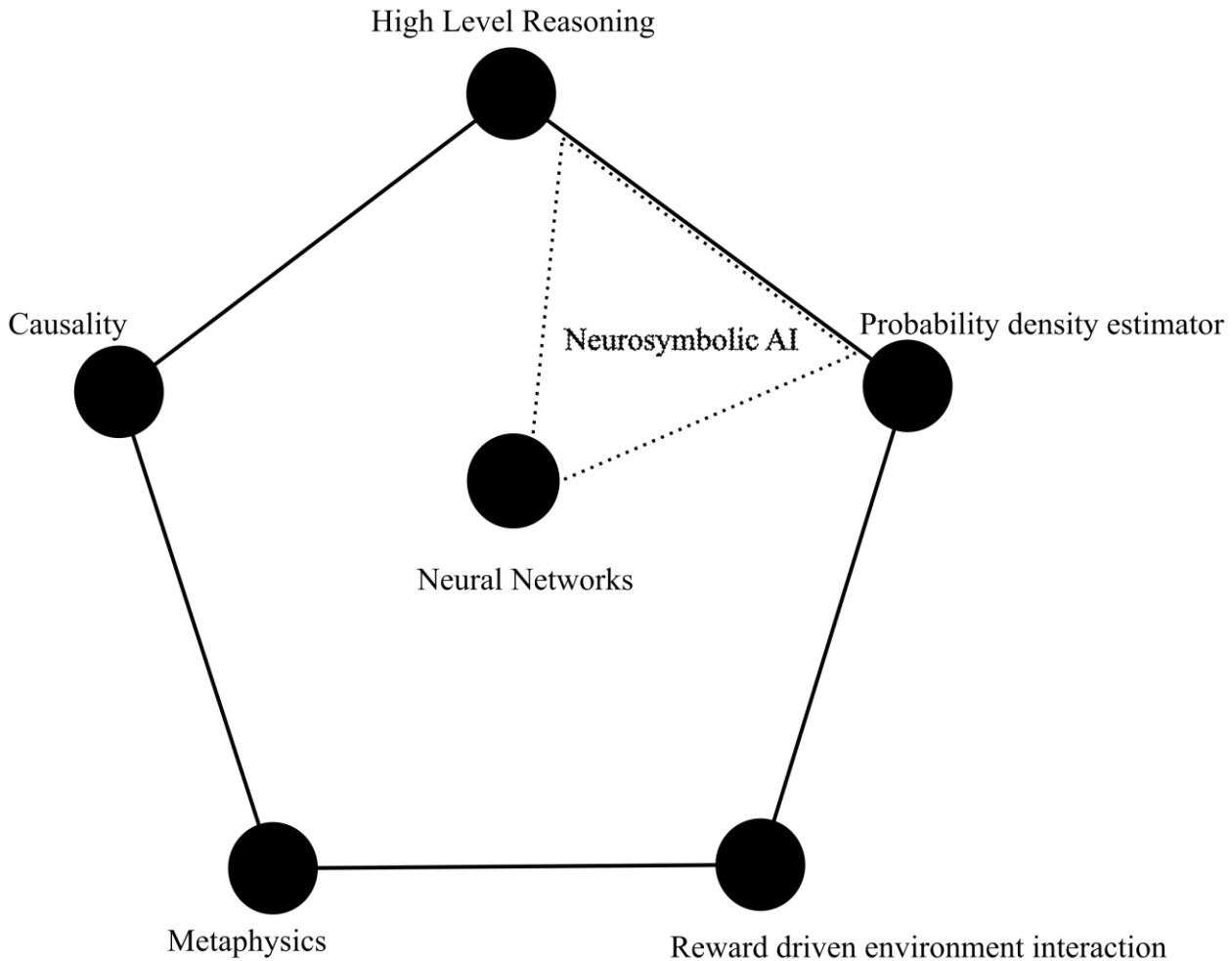


Charting the Landscape of Neuro-Symbolic Reasoners

An Xuelong, supervised by Dr. Antonio Vergari

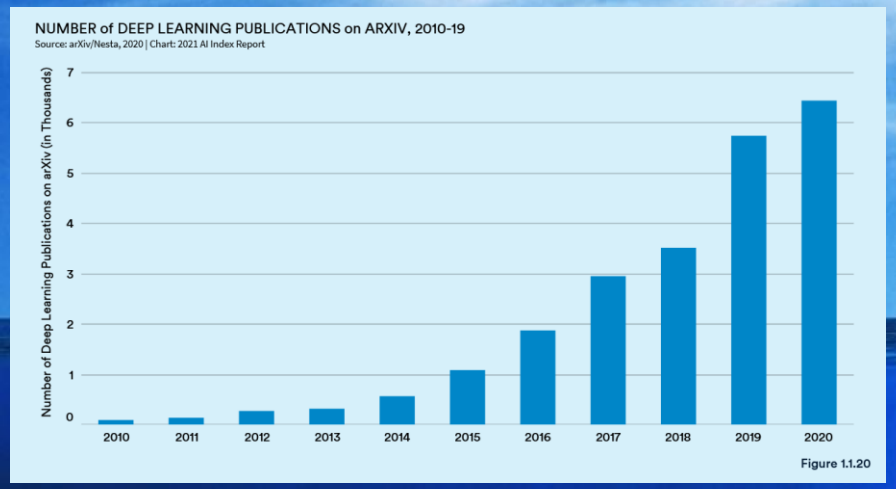
Outline

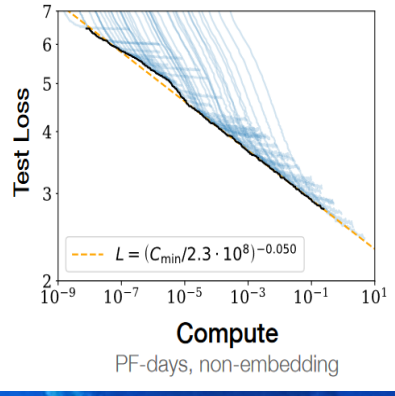
Motivation

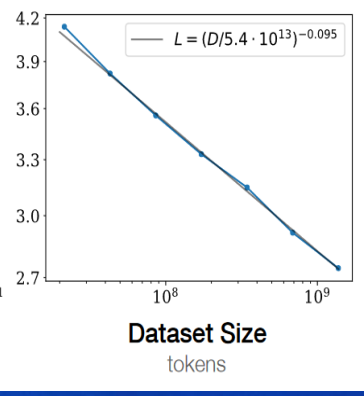
Research question & Contributions

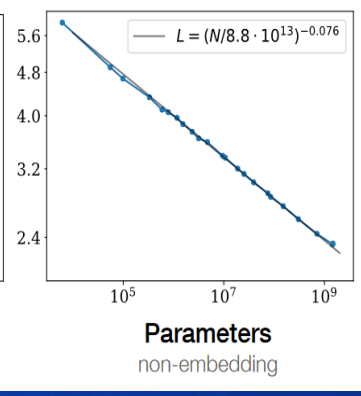
Research Progress

Questions and Feedback

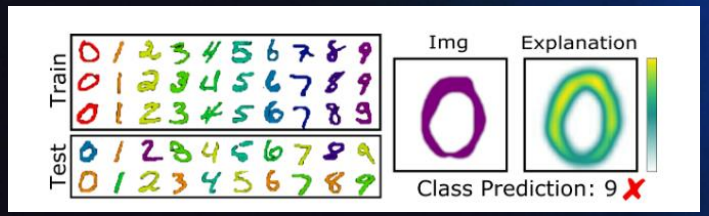

NUMBER of DEEP LEARNING PUBLICATIONS on ARXIV, 2010-19
Source: arXiv/Nesta, 2020 | Chart: 2021 AI Index Report
Number of Deep Learning Publications on arXiv (in Thousands)
0
1
2
3
4
5
6
7
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
Figure 1.1.20

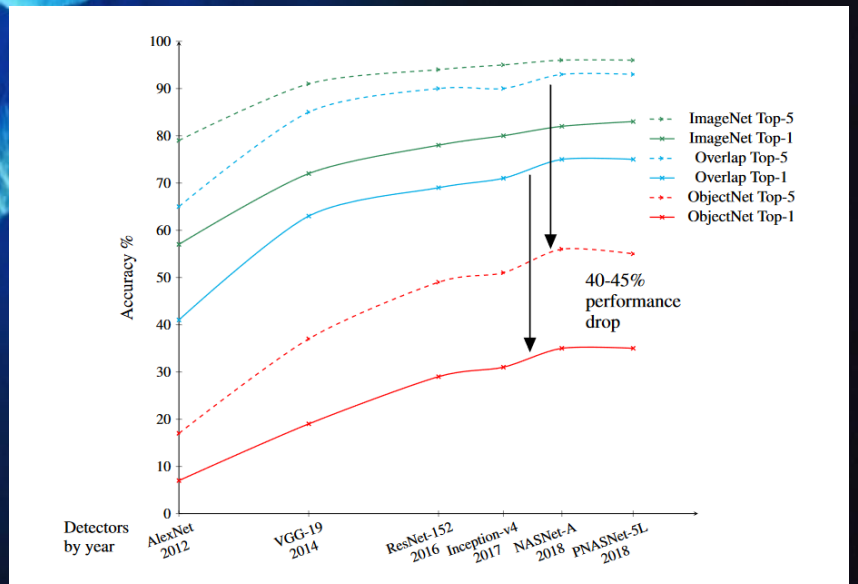

Test Loss
$L = (C_{min}/2.3 \cdot 10^8)^{-0.050}$
Compute
PF-days, non-embedding


$L = (D/5.4 \cdot 10^{13})^{-0.095}$
Dataset Size
tokens

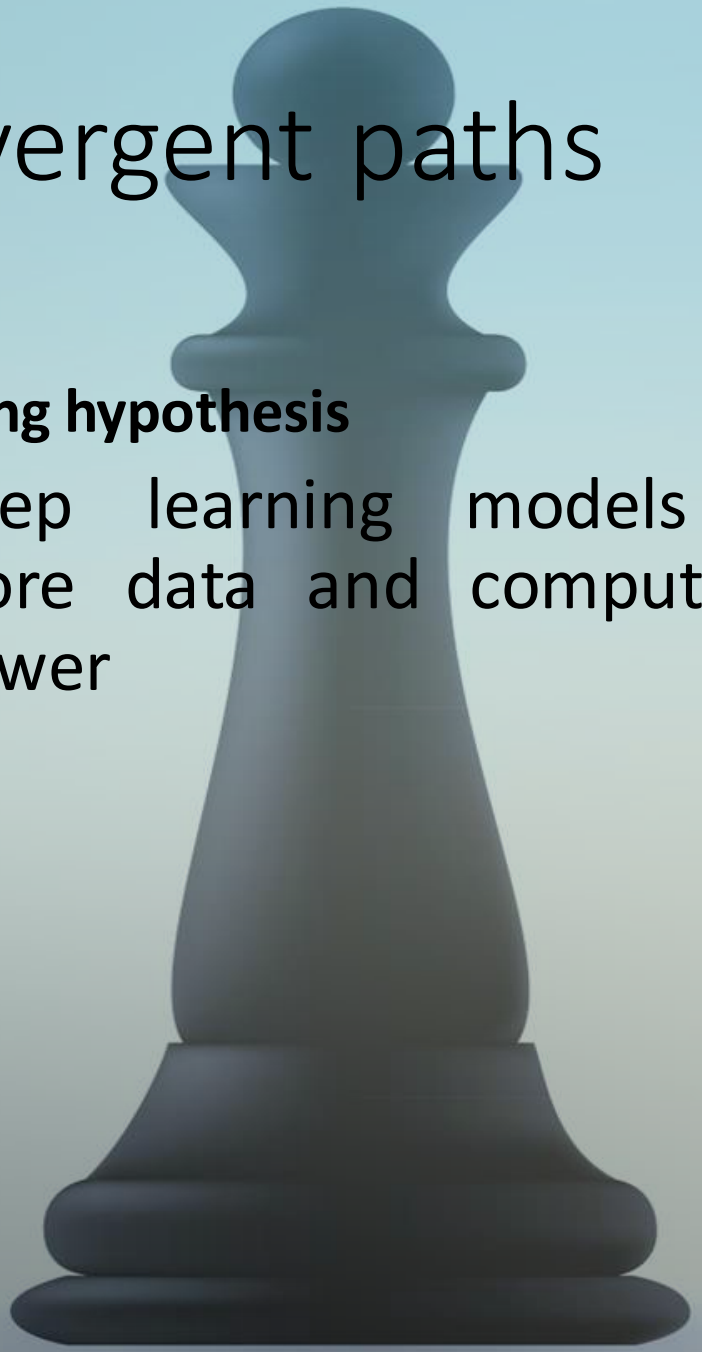

$L = (N/8.8 \cdot 10^{13})^{-0.076}$
Parameters
non-embedding


ImageNet
Chairs
ObjectNet
Chairs by rotation
Chairs by background
Chairs by viewpoint
Teapots
T-shirts


Train
Test
Img
Explanation
Class Prediction: 9


Accuracy %
ImageNet Top-5
ImageNet Top-1
Overlap Top-5
Overlap Top-1
ObjectNet Top-5
ObjectNet Top-1
40-45% performance drop
Detectors by year
AlexNet 2012
VGG-19 2014
ResNet-152 2016
Inception-v4 2017
NASNet-A 2018
PNASNet-5L 2018

Divergent paths



Scaling hypothesis

- Deep learning models need more data and computational power

Integration hypothesis

- School of symbolism and connectionism have properties that can complement each other's faults

Complementarity vs. dichotomy

Emphysema prediction

a) **Logic Program**

```
friendOf(kerry,  natalia).
friendOf(natalia, shirley).
smoker(natalia).
emphysema(natalia).
smoker(X) :- friendOf(X, Y),
             smoker(Y)
emphysema(X) :- smoker(X)

query(emphysema(kerry), ?)
```

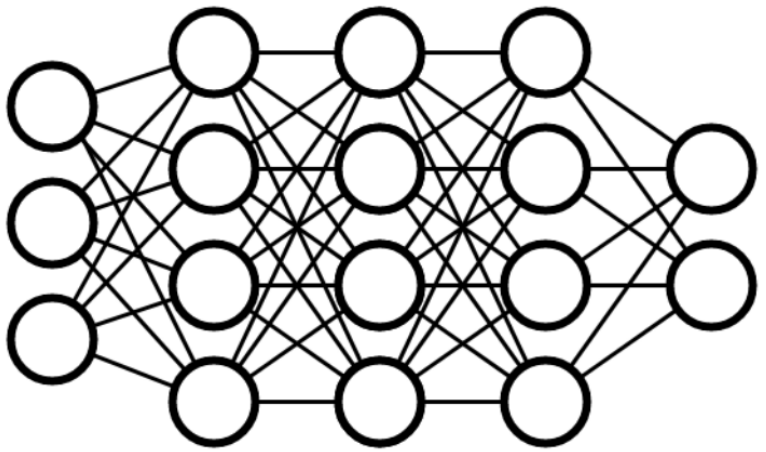
Transparent

Parsimonious

Inflexible

b) **Deep Neural Network**

x1	x2	x3	x4	x5	x6	x7	y
35,5	no	yes	no	no	no	no	no
35,9	no	no	yes	yes	yes	yes	no
35,9	no	yes	no	no	no	no	no
36,0	no	no	yes	yes	yes	yes	no
36,0	no	yes	no	no	no	no	no
36,0	no	yes	no	no	no	no	no
36,2	no	no	yes	yes	yes	yes	no
36,2	no	yes	no	no	no	no	no
36,3	no	no	yes	yes	yes	yes	no
36,6	no	no	yes	yes	yes	yes	no



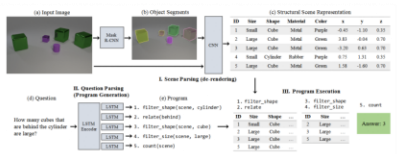
Semi/un-structured data

Black-box

Scalable

Strenghts

Brittleness



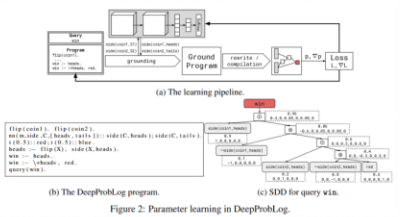
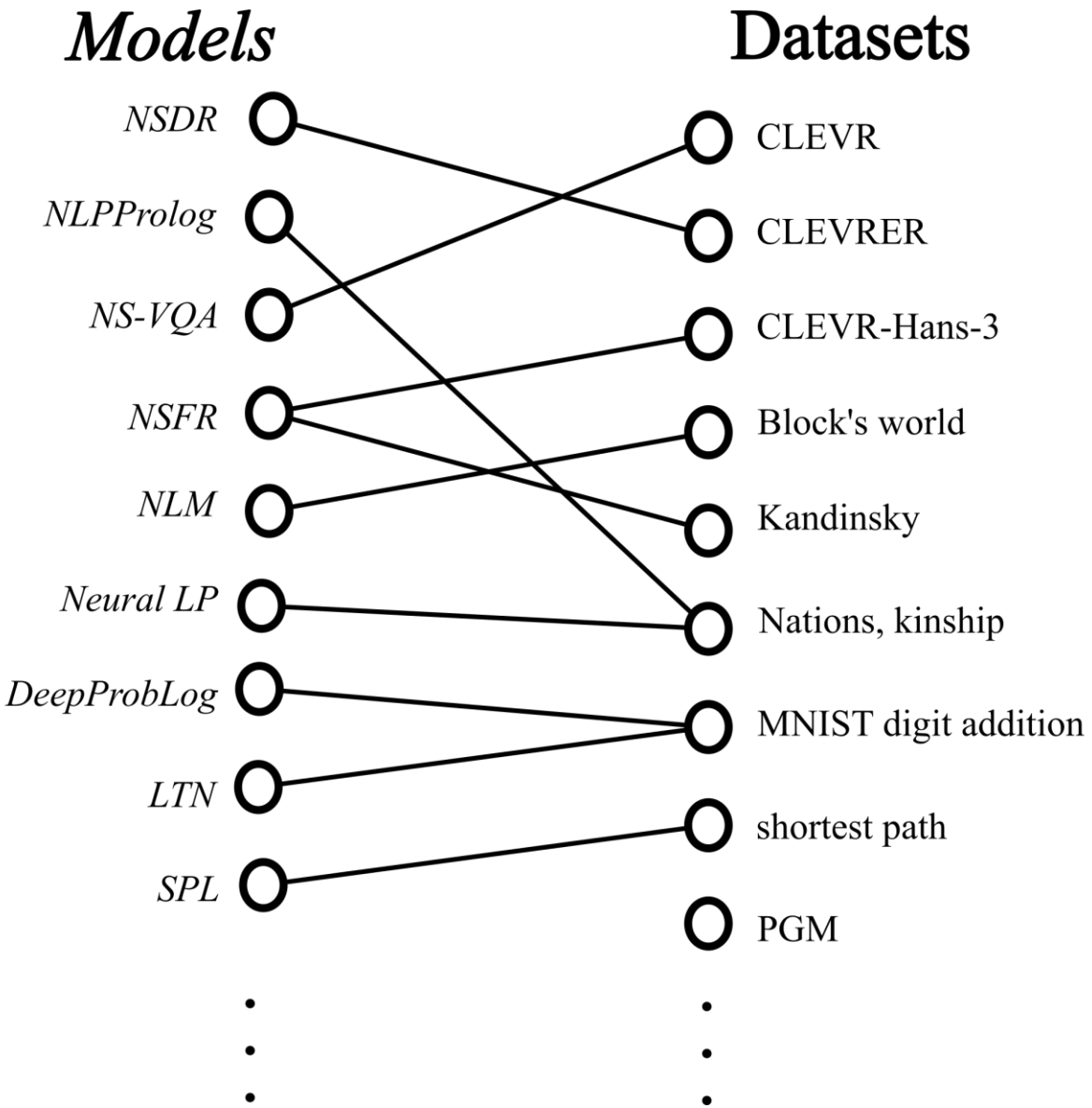
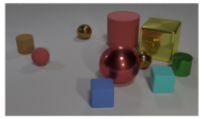
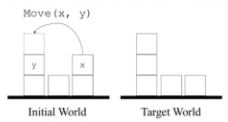
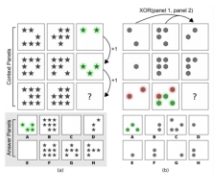


Figure 2: Parameter learning in DeepProbLog.









"NeSy AI is in need of standard benchmarks … [to] provide a fair comparative evaluation of different approaches…(Garcez & Lamb, 2020)

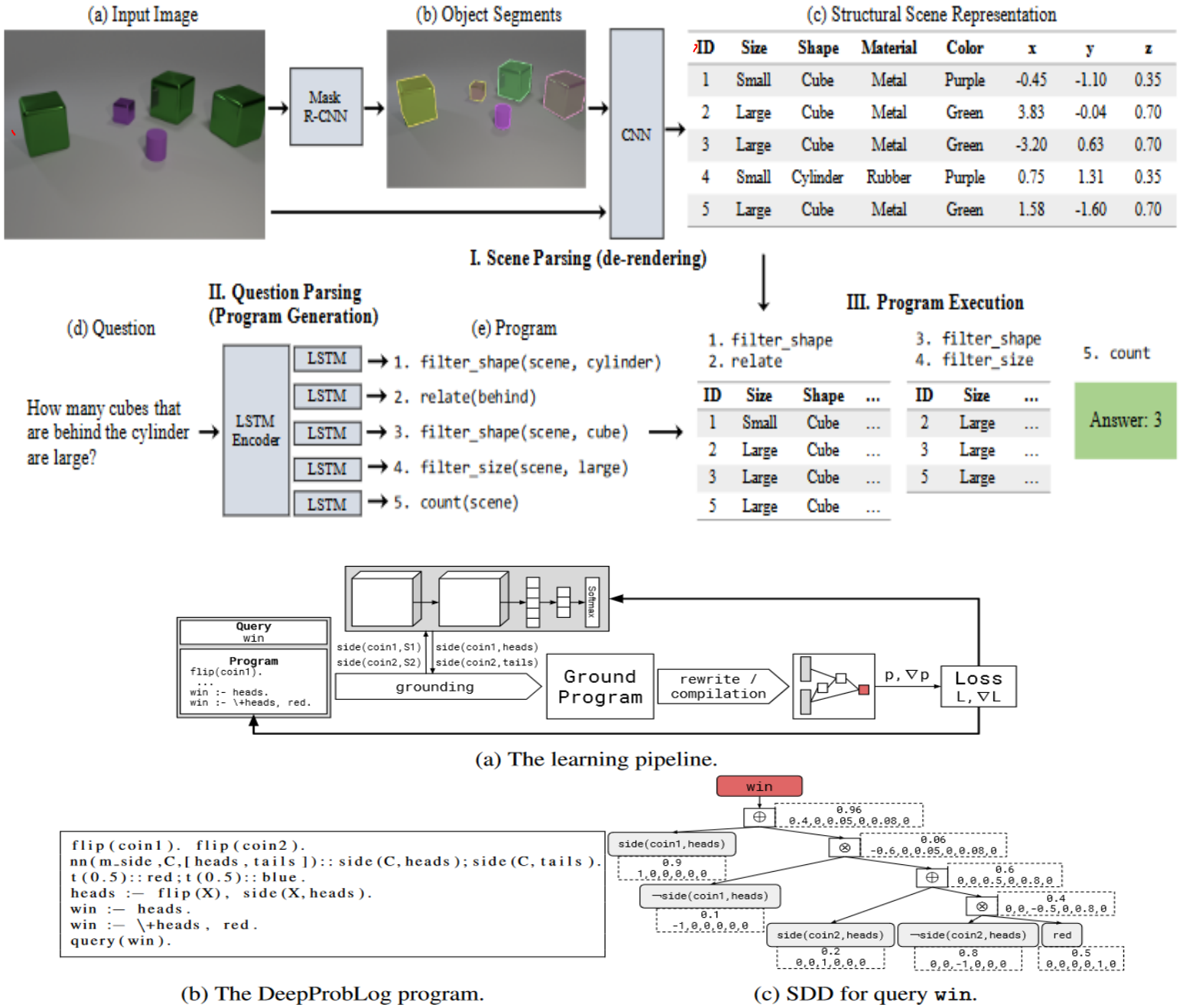
Contributions

Theoretical side:

- Unified framework for NeSy model
- Unified framework for benchmark

Practical side:

- Devise common benchmark for testing current and future models



ID	Size	Shape	Material	Color	x	y	z
1	Small	Cube	Metal	Purple	-0.45	-1.10	0.35
2	Large	Cube	Metal	Green	3.83	-0.04	0.70
3	Large	Cube	Metal	Green	-3.20	0.63	0.70
4	Small	Cylinder	Rubber	Purple	0.75	1.31	0.35
5	Large	Cube	Metal	Green	1.58	-1.60	0.70

ID	Size	Shape	...
1	Small	Cube	...
2	Large	Cube	...
3	Large	Cube	...
5	Large	Cube	...

ID	Size	...
2	Large	...
3	Large	...
5	Large	...

Figure 2: Parameter learning in DeepProbLog.

Yi, K. et. al (2019)

Manhaeve, R. et. al (2018)

Frameworks	Inference	Syntax	Semantics	Learning	Representations	Paradigms	Tasks
	(P)roof (M)odel	(P)ropositional (R)elational (FOL)	(M)inimal (S)table (C)lassical (F)uzzy (P)robability	(P)arameters (S)tructure	(S)ymbolic (Sub)symbolic	Logic (L/l) Probability (P/p) Neural(N/n)	(D)istant (S)upervision (S)emi (S)upervised (KGC)ompletion (G)enerative (K)nowledge (I)nduction
αILP [109]	P+M	FOL	S + P	P + S	S	Ln	KI
∂ILP [38]	P	R	M + F	P + S	S	Ln	DS + KI
DeepProbLog [72]	P+M	FOL	M + P	P+S	S+Sub	LpN	DS + KI
DeepStochLog[129]	P	FOL	M + P	P	S	LpN	DS + SS
DiffLog[110]	P	R	M + F	P+S	S	Ln	KI
DL2[39]	M	P	C + F	P	S+Sub	lN	DS + SS
DLM[77]	M	FOL	C + F + P	P	S	lPN	SS + KGC
LRNN[140]	P	R	M + F	P + S	S + Sub	LN	KGC + KI
LTN[6]	M	FOL	C + F	P	S + Sub	lN	DS + SS
NeuralLP[134]	P	R	M + F	P	S	Ln	KGC + KI
NeurASP[135]	P+M	FOL	S + P	P	S	LpN	DS
NLM[34]	P	R	M + F	P + S	S	Ln	KGC + KI
NLog[118]	P	R	M + P	P	S	LpN	DS
NLProlog[128]	P	R	M + P	P + S	S + Sub	LpN	KGC + KI
NMLN[78]	M	FOL	C + P	P + S	S + Sub	lPN	KGC + G
NTP[100]	P	R	M + F	P + S	S + Sub	Ln	KGC + KI
RNM[76]	M	FOL	C + P	P	S	lPN	SS
SBR[32]	M	FOL	C + F	P	S+Sub	lN	DS + SS
Scallop[59]	P	FOL	M + P	P	S	LpN	DS
SL[130]	M	P	C + P	S	S	LpN	SS
Slash[111]	P+M	FOL	S + P	P	S	LpN	DS +SS
TensorLog[18]	P	R	M + P	P	S	LpN	DS + KGC

Table 1: Logic-based NeSy frameworks according to the 6 dimensions outlined in the paper.

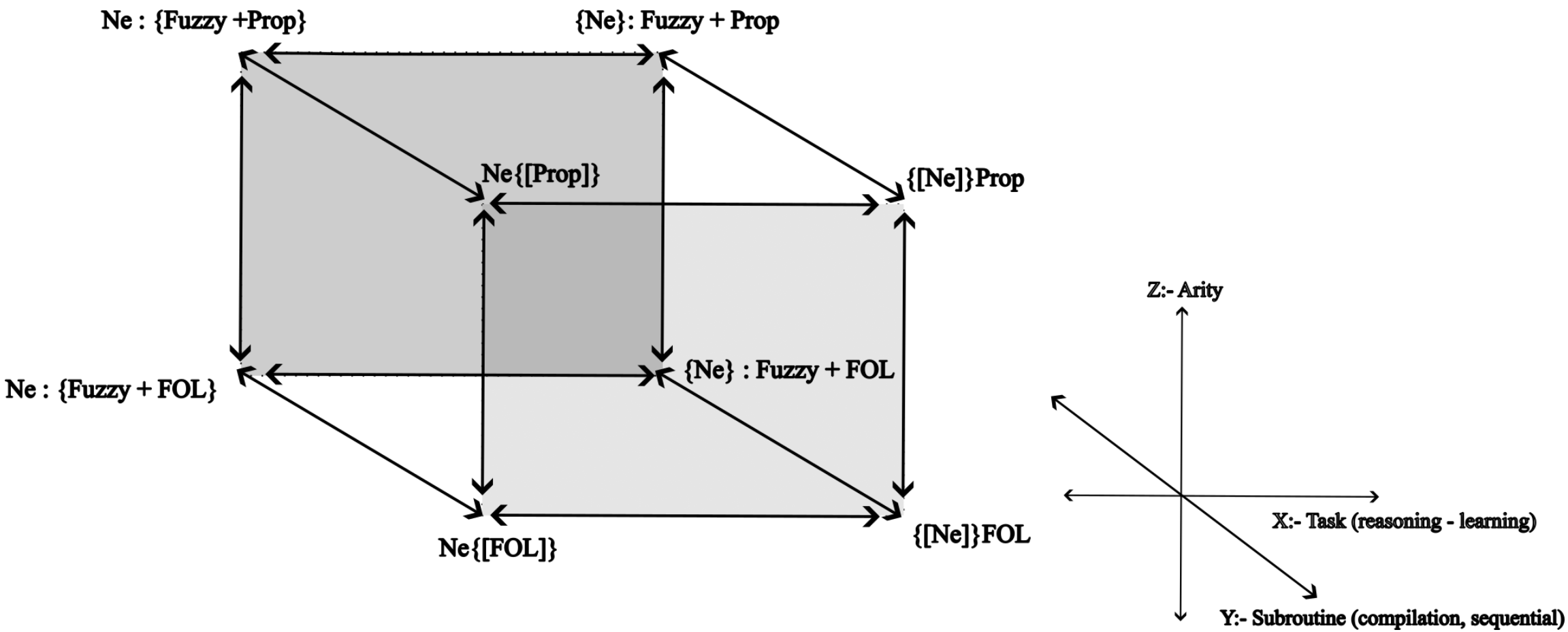
Luc De Raedt et al., 2020 (revised on May 21, 2023)

category	number of papers	papers
[symbolic Neuro symbolic]	13	[9–17]
[Symbolic[Neuro]]	9	[15, 18–23]
[Neuro ∪ compile(Symbolic)]	10	[24–33]
[Neuro → Symbolic]	13	[8, 23, 34–44]
[Neuro[Symbolic]]	0	N/A

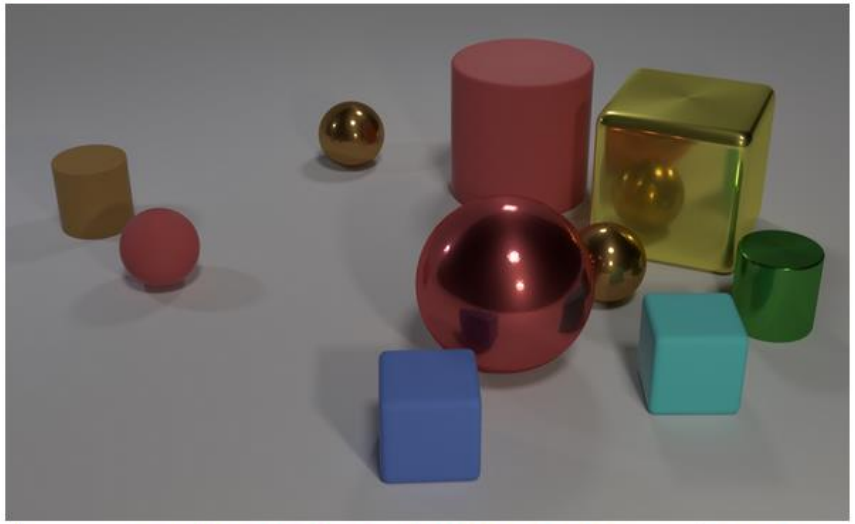
Table 2

Kautz categories paper count. Two paper fit two categories.

Md K. Saker et al., 2021


Ne : {Fuzzy +Prop}
{Ne}: Fuzzy + Prop
Ne{[Prop]}
{[Ne]}Prop
Ne : {Fuzzy + FOL}
{Ne} : Fuzzy + FOL
Ne{[FOL]}
{[Ne]}FOL
Z:- Arity
X:- Task (reasoning - learning)
Y:- Subroutine (compilation, sequential)

Questions in CLEVR test various aspects of visual reasoning including **attribute identification**, **counting**, **comparison**, **spatial relationships**, and **logical operations**.



Q: Are there an **equal number** of **large things** and **metal spheres**?

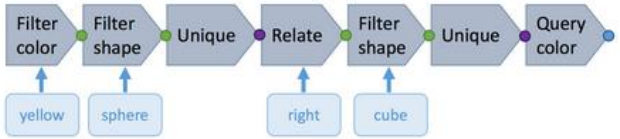
Q: **What size** is the **cylinder** **that is left of** the **brown metal** thing **that is left of** the **big sphere**?

Q: There is a **sphere** with the **same size as** the **metal cube**; is it **made of the same material as** the **small red sphere**?

Q: **How many** objects are **either** **small cylinders** or **red** things?

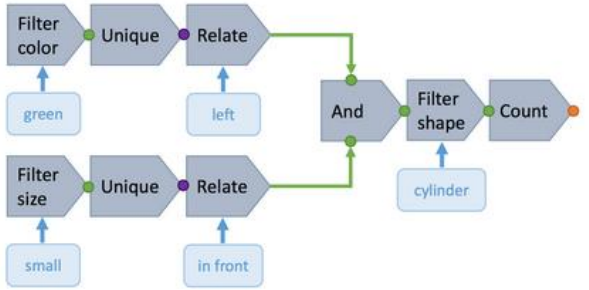
Each question in CLEVR is represented both in **natural language** and as a **functional program**. The functional program representation allows for precise determination of the reasoning skills required to answer each question.

Sample chain-structured question:



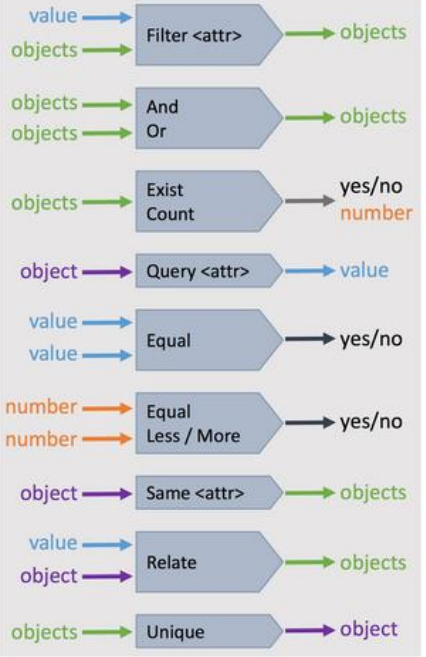
What color is the cube to the right of the yellow sphere?

Sample tree-structured question:



How many cylinders are in front of the tiny thing and on the left side of the green object?

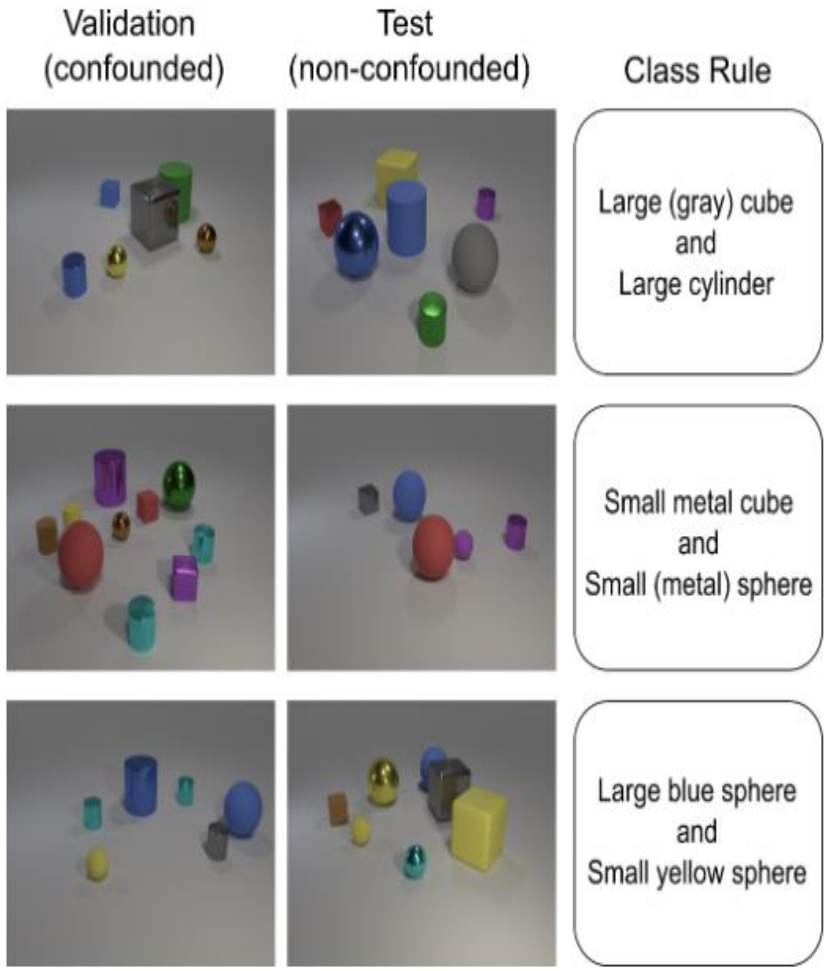
CLEVR function catalog



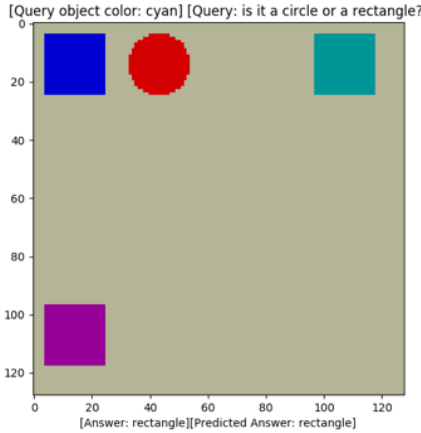
- Exist, Count, Compare Integer, Query Attribute, and Compare Attribute,

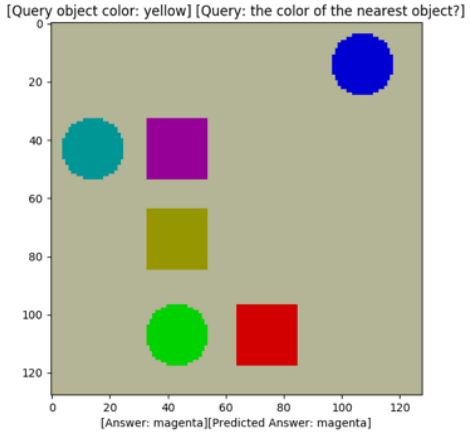
Johnson, J. et. al (2017)

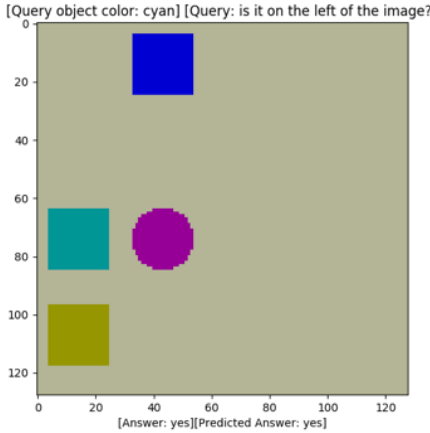
CLEVR-Hans3 (Stammer, et. Al 2021)



Sort-of-CLEVR (Santoro, A., et. Al (2017))

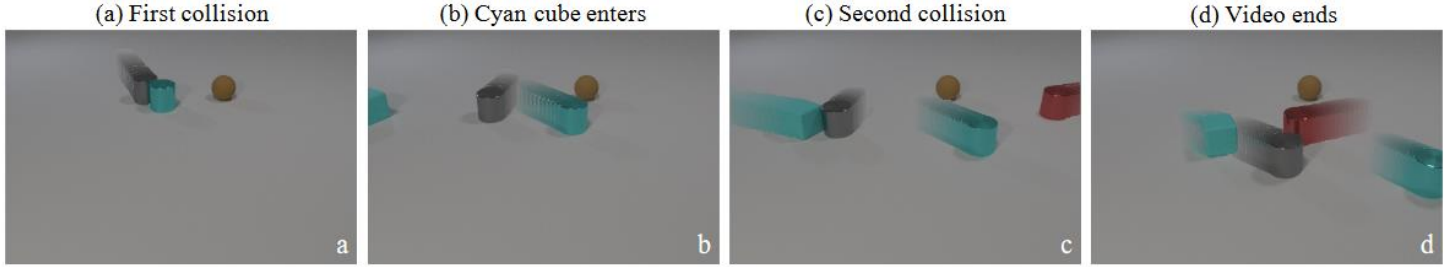






- Query Attribute, and Compare Attribute,

CLEVRER (Kexin, Yi, et. al, 2020)



I. Descriptive

Q: *What shape is the object that collides with the cyan cylinder?* **A:** *cylinder*

Q: *How many metal objects are moving when the video ends?* **A:** *3*

II. Explanatory

Q: *Which of the following is responsible for the gray cylinder's colliding with the cube?*

a) The presence of the sphere

b) The collision between the gray cylinder and the cyan cylinder **A:** *b)*

III. Predictive

Q: *Which event will happen next*

a) The cube collides with the red object

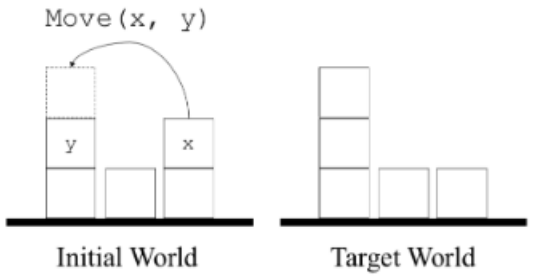
b) The cyan cylinder collides with the red object **A:** *a)*

IV. Counterfactual

Q: *Without the gray object, which event will not happen?*

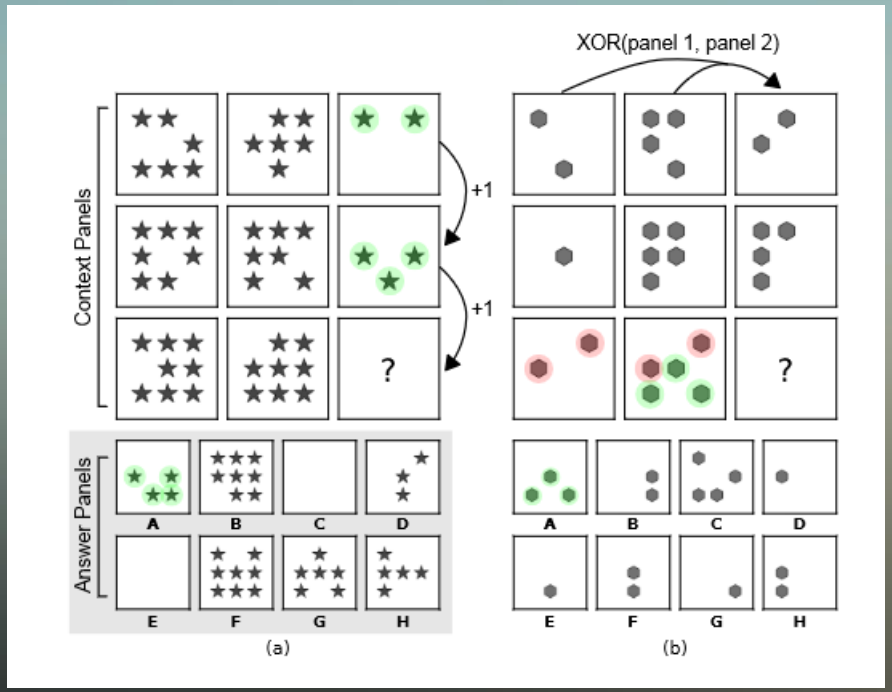
a) The cyan cylinder collides with the sphere

b) The red object and the sphere collide **A:** *a), b)*



$\texttt{On}(x, y)$	True if x is on y
$\texttt{IsGround}(x)$	True if x is the ground
$\texttt{Clear}(x)$	True if there is no block on x
$\texttt{Moveable}(x)$	$\neg\texttt{IsGround}(x) \wedge \texttt{Clear}(x)$
$\texttt{Placeable}(x)$	$\texttt{IsGround}(x) \vee \texttt{Clear}(x)$

Dong, H. et. al (2019)



Barret, D. et. al (2018)

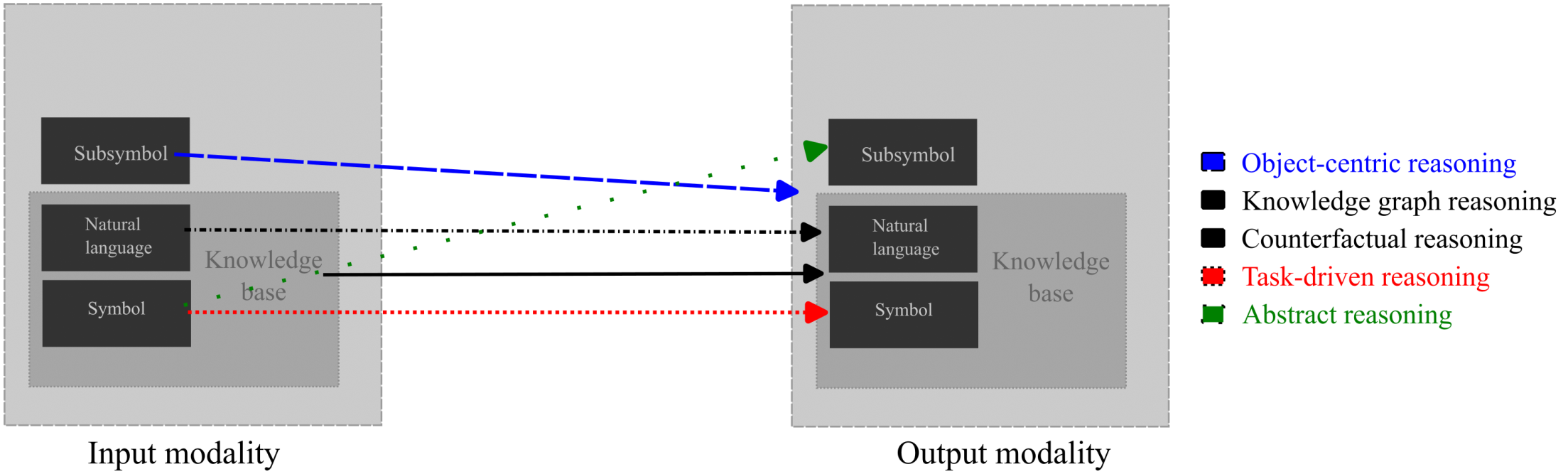
The Hanging Gardens, in **[Mumbai]**, also known as Pherozeshah Mehta Gardens, are terraced gardens … They provide sunset views over the **[Arabian Sea]** …

Mumbai (also known as Bombay, the official name until 1995) is the capital city of the Indian state of Maharashtra. It is the most populous city in **India** …

The **Arabian Sea** is a region of the northern Indian Ocean bounded on the north by **Pakistan** and **Iran**, on the west by northeastern **Somalia** and the Arabian Peninsula, and on the east by **India** …

Q: (Hanging gardens of Mumbai, country, ?)
Options: {Iran, **India**, Pakistan, Somalia, …}

Dataset Curated vs. Prescribed

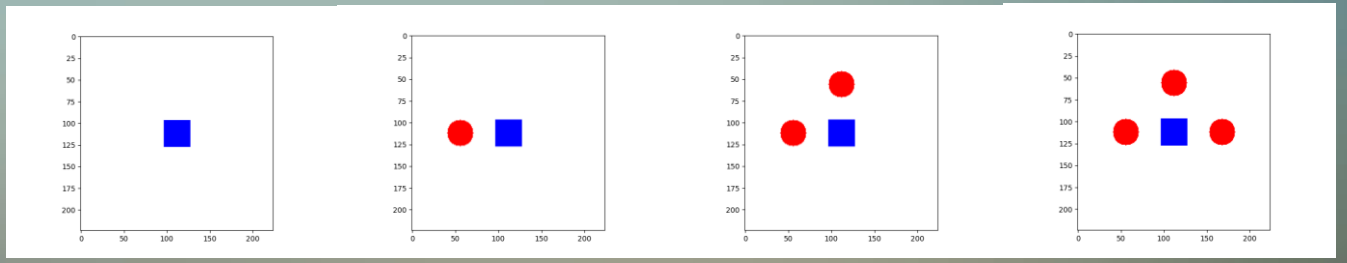


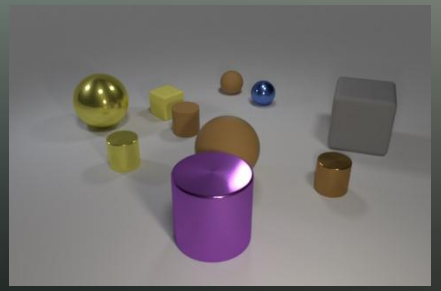
Taxonomy of NeSy benchmarks

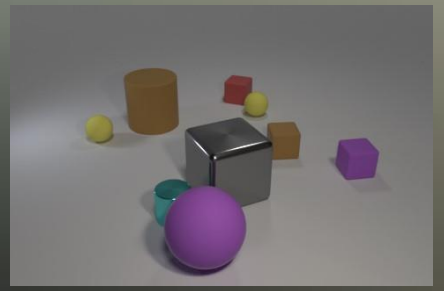
Nature of Task	Format	Input-output during inference	Datasets
Object-centric relational reasoning	Natural language description of scenes with some predicates in .json	Query and image - answer	CLEVR CLEVR-CoGenT CLEVR-Hans Kadinsky Patterns
Task-driven reasoning	Specify a set of axioms and the model completes a goal	Axioms-goal completion	Linear regression (real state dataset) Clustering (Badreddine, S. et al., 2021) Block's world problem
Knowledge graph reasoning	NL description of knowledge OR set of facts and relations	Query-Answer	Wiki-hop KB Med-Hop KB
Object centric abstract reasoning	IQ like images and the model must complete the pattern	Image - Image	Procedurally Generated Matrices
Counterfactual reasoning	Hypothetical queries	Video and query - answer	CLEVRER

SaSSy-CLEVR Overview

Object centric relational reasoning	Task driven reasoning	Knowledge graph reasoning	Abstract reasoning	Counterfactual reasoning
CLEVR-Hans-3	Elaborate on CLEVR-Hans3	Synthetize KG from CLEVR-Hans3	Elaborate on Sort-of-CLEVR	Hypothetical queries over CLEVR



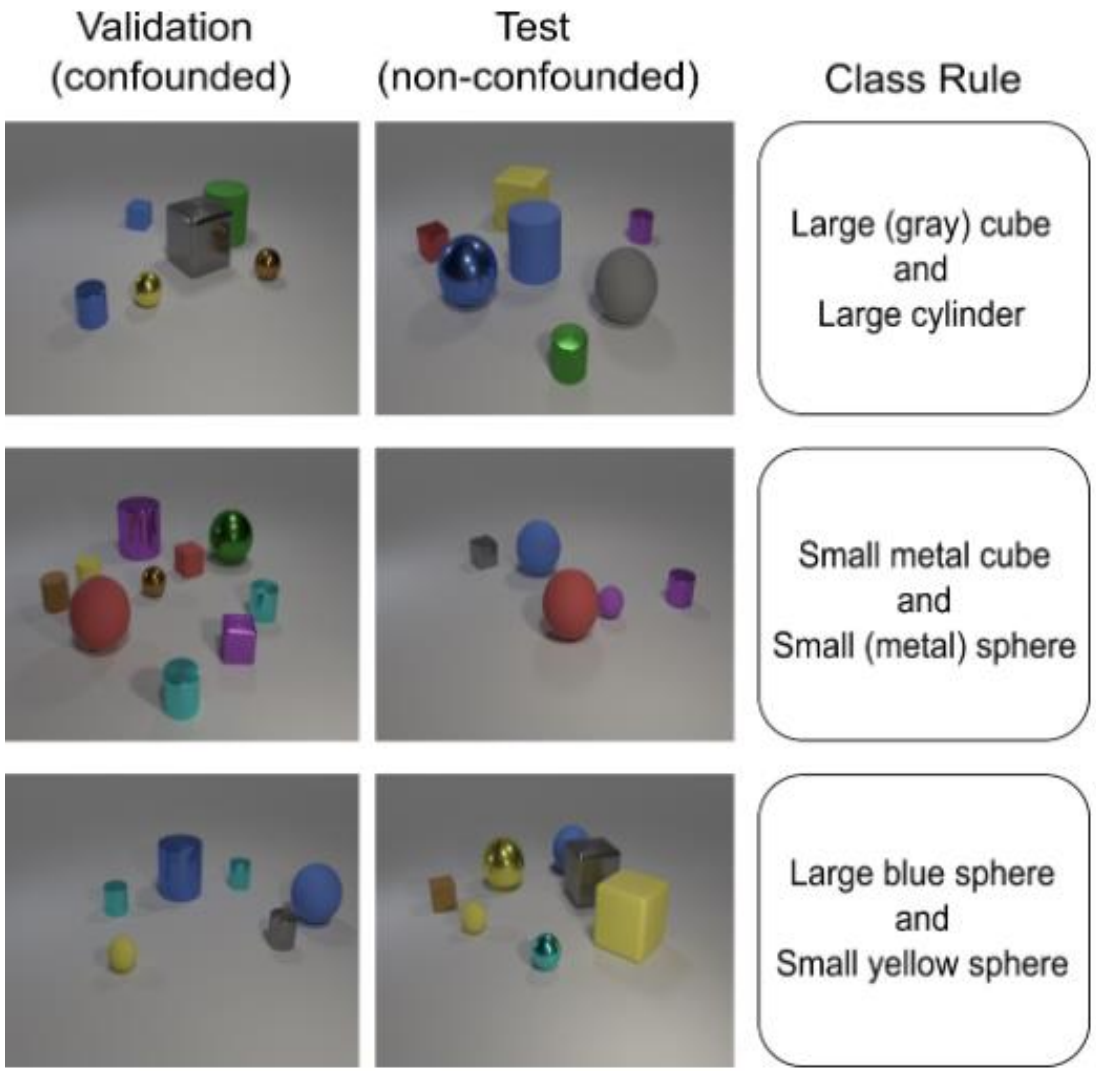




If the blue metal cube is taken, how many objects are left?

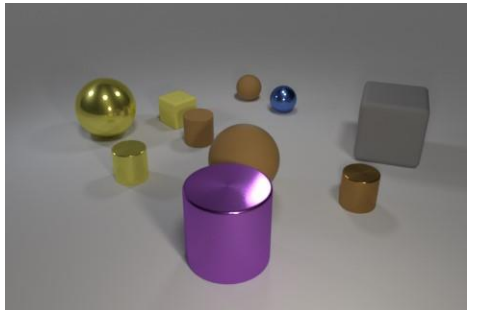
CLEVR-Hans3 for object-centric and KG reasoning

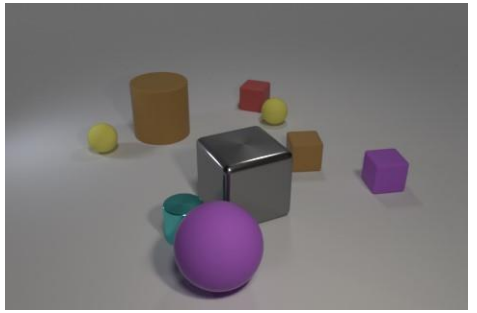
- Visual confounder in test set
- Generate KG



Task-driven reasoning

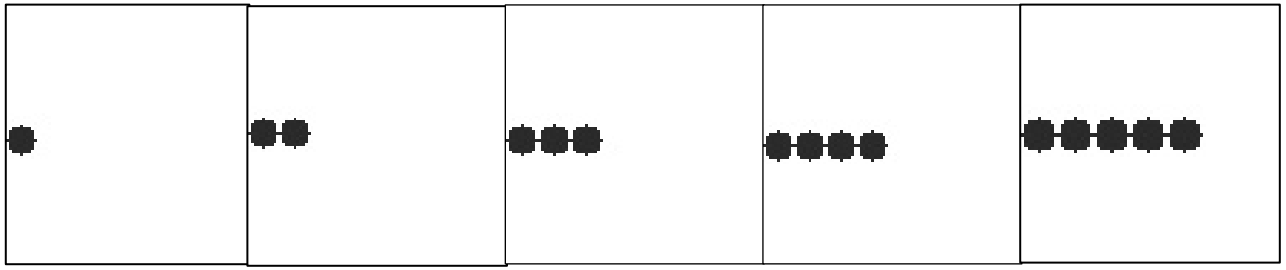
- Possible actions:
  - Move(object, X, Y, Z)
  - Add(object, shape, color, size, color)
  - ChangeColor(object, color)
  - Remove(object)
- Evaluated using Hungarian Loss
- More rules?



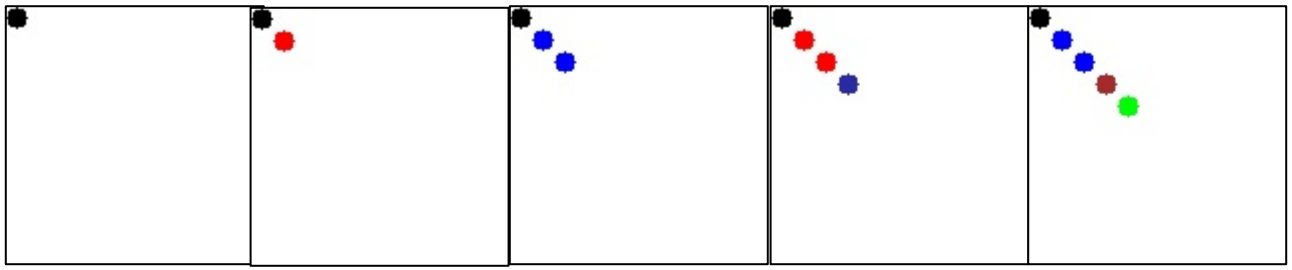


Abstract reasoning: elaborating Sort-of-CLEVR

- Attributes:
  - 2 shapes
  - 7 colours
  - 2 sizes
  - Maximum 10 objects per panel



- 5-panel image sequences: model must complete the 5th panel
  - Arithmetic progression
  - Clockwise movement
  - Alternating



- Visual confounder during testing
  - Colour
  - Shape
  - Rotation?

Counterfactual reasoning

- Query Attribute: What color is the thing right of the red sphere?
- Counting: How many red cubes are there?
- Existence: Are there any cubes to the right of the red thing?
- Compare Integer: Are there fewer cubes than red things?

- Query Attribute: What color is the thing right of the red sphere if the blue cube is removed?
- Counting: how many objects will there be if the blue metal cube is removed?
- Existence: Will there be any cubes to the right of the red thing if the blue cube is removed?
- Compare Integer: Will there be fewer cubes than red things if the red cube is removed?

Appendix

Catalogue of Benchmarks based on Five Major Reasoning Tasks

Nature of Task	Input-Output	Challenging Aspects	Examples
Object-Centric Relational Reasoning	Images and Query - Answer	Confounding concepts; out-of-distribution generalization; interpretability	CLEVR, Kandinsky Patterns, CLEVR-Hans
Multi-Hop Reasoning	KB and Query - Answer	Satisfiability; search efficiency	ChEMBL, Wiki-Hop, MedHop, babi
Task-Driven Reasoning	Logic program and Query - Answer	Satisfiability; search efficiency	Block's World, Sorting Arrays, Coin Ball
Object-Centric Abstract Reasoning	Image - Image	Patterns in images are implicit; interpretability	PGM
Counterfactual Reasoning	Image/KB and Query - Answer	Objects in the query are absent in dataset; interpretability	CLEVRER

NeSy models	Neural components	Symbolic components	Categorization
DreamCoder	Program recognition module	Program synthesis	{[Ne]}HOL or {[Ne]}{[HOL]}
NeSy-Visual Question Answering (VQA)	Mask RCNN for image recognition + LSTM to parse questions	SQL-like Query executor	{[Ne]}HOL
αILP	Pretrained Slot Attention	Differentiable forward reasoner	Ne{[FOL]}
DeepProbLog	User-specified neural network	ProbLog using probabilistic circuits for scalable inference	Ne{[FOL]} or {[Ne]}{[FOL]}
Neural Markov Logic Networks (NMLM)	Standard feedforward neural networks to represent factor graph	Probabilistic inference	Ne{[FOL]}
Logic Tensor Networks	User-specified neural network	User-specified logic program in Real Logic	Ne : {Fuzzy FOL}
Relational Neural Machine	Neural Markov Logic Networks	Weighted probabilistic inference	Ne : {Fuzzy FOL}