From Statistical Relational to Neurosymbolic AI

Giuseppe Marra

(joint work with **Luc De Raedt** and many other people, ack at the end)

DTAI
MACHINE LEARNING

KU LEUVEN

LEUVEN.AI INSTITUTE

The neurosymbolic integration quest

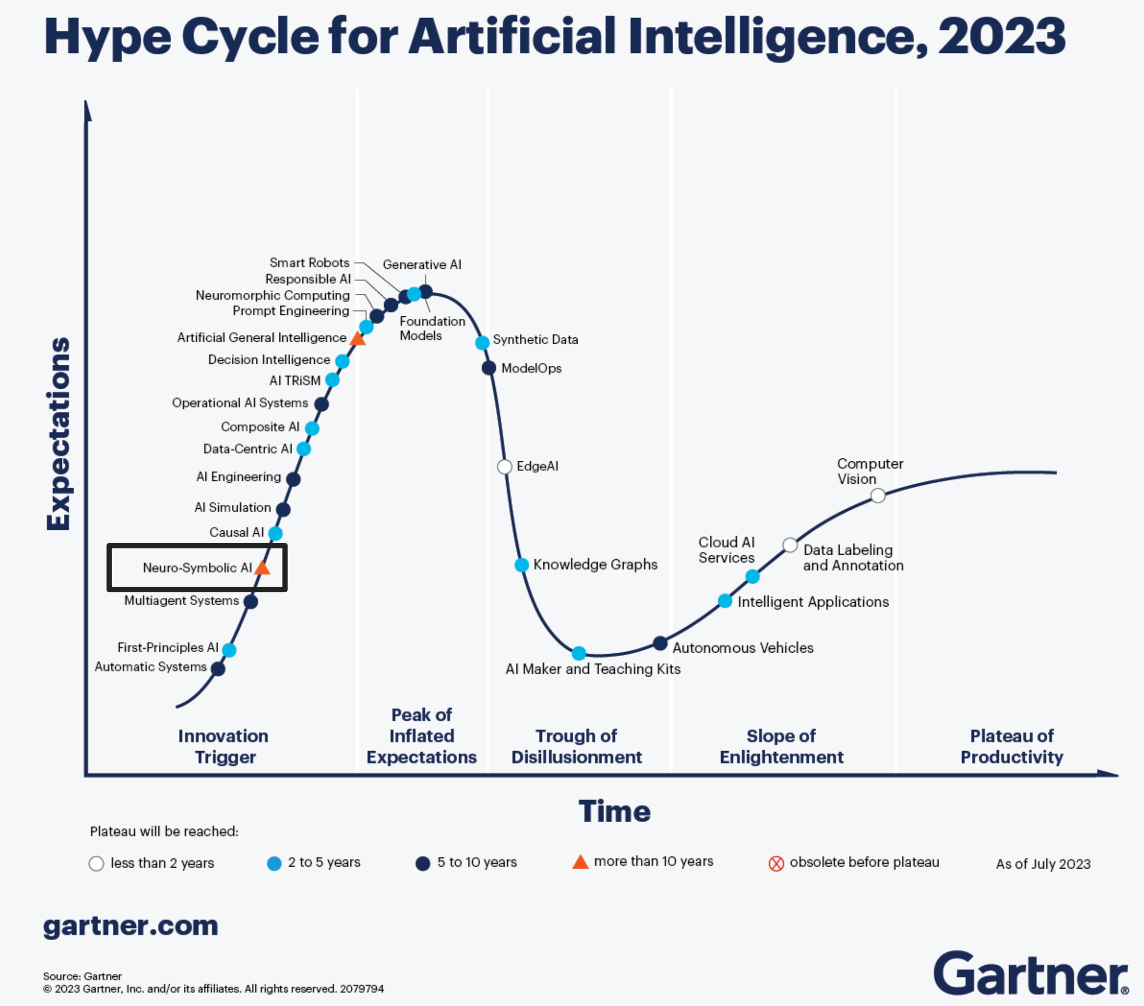
Subsymbolic Approaches	Symbolic Approaches
associative	logic
data	knowledge
learning	reasoning
noisy input	precise input

"We need System 2 Deep Learning " *by Y. Bengio - NeurIPS 2019*

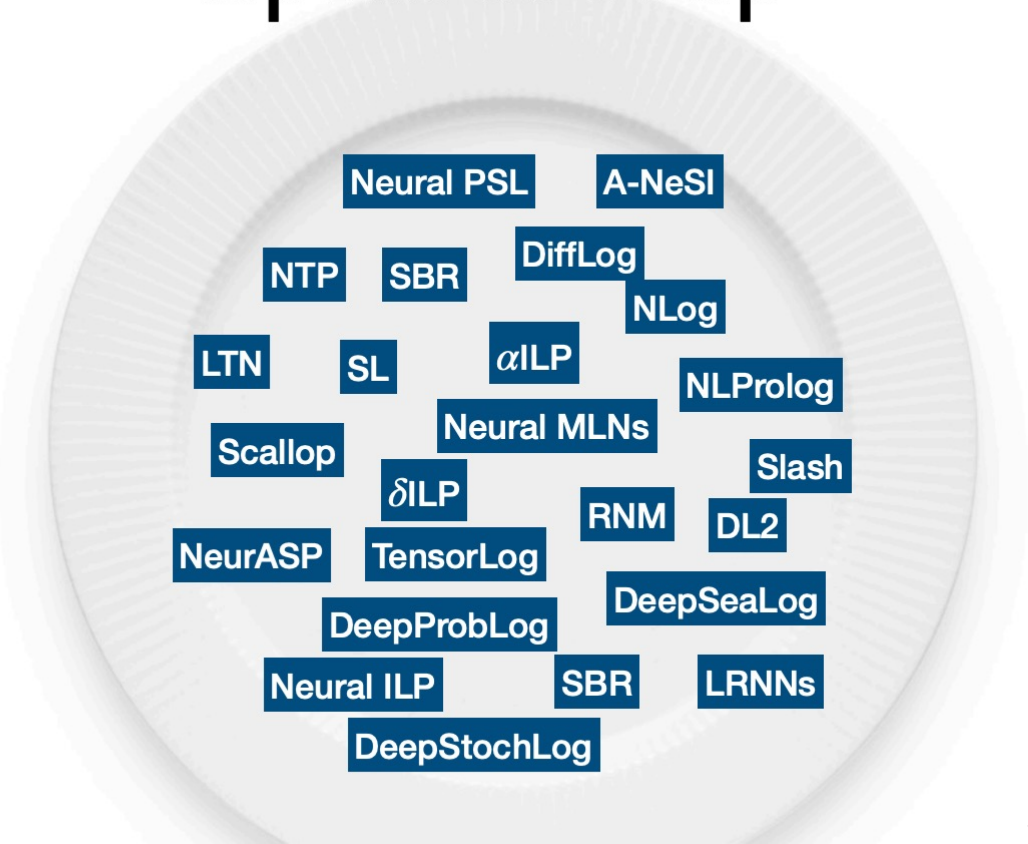
"We need to bring together the neural and symbolic traditions " *by H. Kautz - AAAI 2020*

"The most promising approach to a broad AI is a neuro-symbolic AI, that is, a bilateral *AI* that combines methods from symbolic and sub-symbolic AI." *by Sepp Hochreiter Communications of the ACM, April 2022*

"AlphaGeometry is a neuro-symbolic system made up of a neural language model and a symbolic deduction engine" *by deepmind.google, January 2024*



The NeuroSymbolic alphabet-soup



From Statistical Relational to Neurosymbolic AI

- Another paradigm for learning and reasoning

StarAI = Logic + PGMs

From Statistical Relational to Neurosymbolic AI

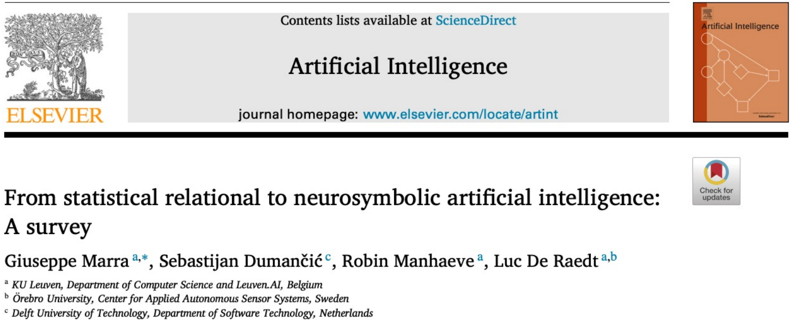
- Another paradigm for learning and reasoning

StarAI = Logic + PGMs

- Can we use insights from this area?

NeSy = Logic + PGMs + Neural Networks

From Statistical Relational to Neurosymbolic AI


Contents lists available at ScienceDirect

Artificial Intelligence

journal homepage: www.elsevier.com/locate/artint

From statistical relational to neurosymbolic artificial intelligence: A survey

Giuseppe Marra[a,*], Sebastijan Dumančić[c], Robin Manhaeve[a], Luc De Raedt[a,b]

[a] KU Leuven, Department of Computer Science and Leuven.AI, Belgium
[b] Örebro University, Center for Applied Autonomous Sensor Systems, Sweden
[c] Delft University of Technology, Department of Software Technology, Netherlands

7 dimensions:
- Proof vs Models
- Logic Syntax
- Logic Semantics
- Structure vs Parameter Learning
- Representations (symbolic vs subsymbolic)
- Recovery of original paradigms
- Tasks

Marra et al, 2024 AIJ
De Raedt et al, 2020, IJCAI

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 [111]	P+M	FOL	S + P	P + S	S	Ln	KI
∂ILP [39]	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 [132]	P	FOL	M + P	P	S	LpN	DS + SS
DiffLog [112]	P	R	M + F	P+S	S	Ln	KI
DL2 [40]	M	P	C + F	P	S+Sub	lN	DS + SS
DLM [77]	M	FOL	C + F + P	P	S	lPN	SS + KGC
LRNN [116]	P	R	M + F	P + S	S + Sub	LN	KGC + KI
LTN [5]	M	FOL	C + F	P	S + Sub	lN	DS + SS
NeuralLP [137]	P	R	M + F	P	S	Ln	KGC + KI
NeurASP [138]	P+M	FOL	S + P	P	S	LpN	DS
NLM [35]	P	R	M + F	P + S	S	Ln	KGC + KI
NLog [121]	P	R	M + P	P	S	LpN	DS
NLProlog [131]	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 [102]	P	R	M + F	P + S	S + Sub	Ln	KGC + KI
RNM [76]	M	FOL	C + P	P	S	lPN	SS
SBR [33]	M	FOL	C + F	P	S+Sub	lN	DS + SS
Scallop [58]	P	FOL	M + P	P	S	LpN	DS
SL [133]	M	P	C + P	P	S	LpN	SS
Slash [113]	P+M	FOL	S + P	P	S	LpN	DS +SS

From Statistical Relational to Neurosymbolic AI

7 dimensions:

- Proof vs Models

- Logic Syntax
- Logic Semantics
- Structure vs Parameter Learning
- Representations (symbolic vs subsymbolic)
- Recovery of original paradigms
- Tasks

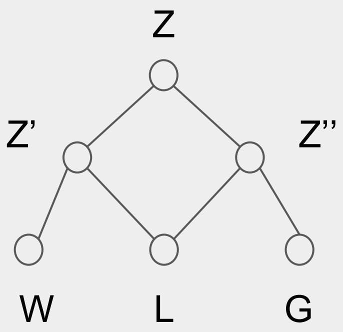
Marra et al, 2024 AIJ
De Raedt et al, 2020, IJCAI

The use of logic: Proof vs Model

The use of logic: Proof vs Model
Logic

Logic rules as
computational rules
(logic programs)

```
Lion,Wall -> Zoo
Lion,Gate -> Zoo.
```



proof-based

Logic rules as
constraints
(SAT)

```
Lion -> Zoo XOR Savanna
```

```
L=?, Z=?, S=?
```

L=T, Z=T, S=F

model-based

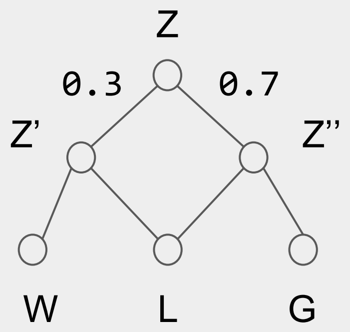
The use of logic: Proof vs Model

StarAI = Logic + Probabilities

Stochastic Logic Programs

(directed graphical models)

```
0.3::Lion,Wall -> Zoo
0.7::Lion,Gate -> Zoo.
```



proof-based

Markov Logic

(undirected graphical models)

```
2.75:: L -> Z XOR S

p(L=T, Z=T, S=F) = ?
```

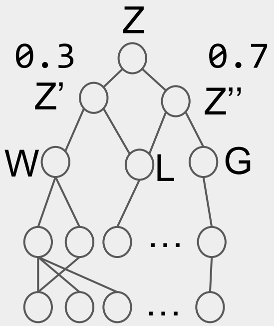
model-based

The use of logic: Proof vs Model

NeSy = Logic + Probabilities + Neural

Logic Programs as
layers/architectures

```
0.3::Lion,Wall -> Zoo
0.7::Lion,Gate -> Zoo.
```



proof-based

Logic-based
Loss Functions /
Energy Function

```
2.75:: L -> Z XOR S

Energy = 2.75 * Satisfaction

e.g.
Loss = CrossEntropy + Energy
```

model-based

From Statistical Relational to Neurosymbolic AI

Can we start from one (StarAI) and build the other (NeSy)?

… a StarAI **recipe** for NeSy …

The use of logic: Model

Semantic Based Regularization

L -> **Z** XOR **S**

G L S Z W

Diligenti et al, AIJ 2017
Marra et al., ECML 2019

The use of logic: Model

Semantic Based Regularization

```
L -> Z XOR S
```

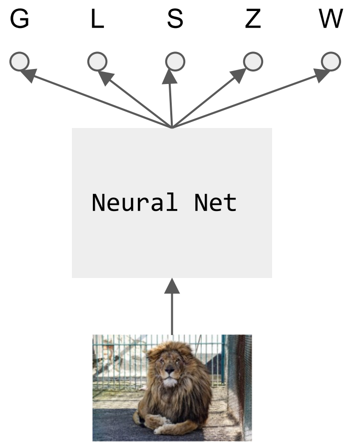
G L S Z W

Diligenti et al, AIJ 2017
Marra et al., ECML 2019

The use of logic: Model

Semantic Based Regularization

The net uses subsymbolic information to find a model.

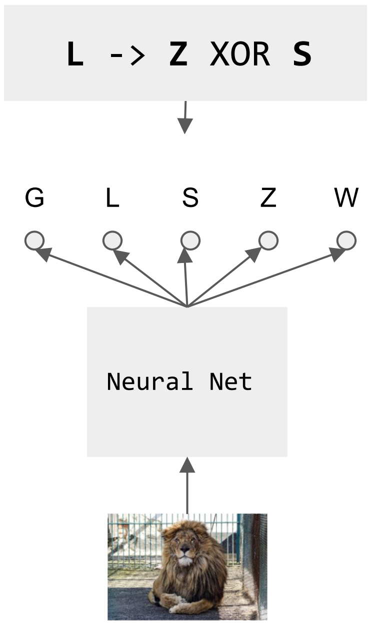


Diligenti et al, AIJ 2017
Marra et al., ECML 2019

The use of logic: Model

Semantic Based Regularization

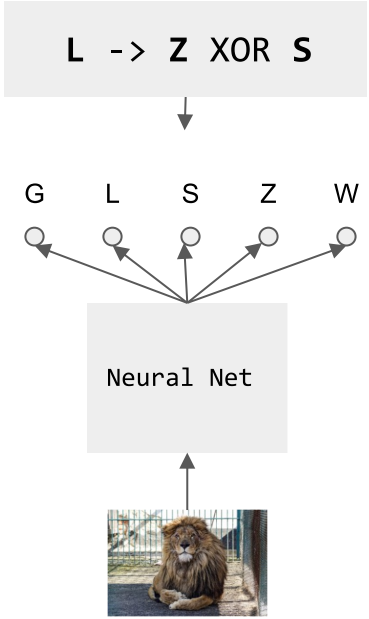
Logic rule is defined on the output variables of the network



Diligenti et al, AIJ 2017
Marra et al., ECML 2019

The use of logic: Model
Semantic Based Regularization

Logic Loss



(Fuzzy) Logic is used as regularization.

TotalLoss = SupervisedLoss + LogicLoss

Diligenti et al, AIJ 2017
Marra et al., ECML 2019

The use of logic: Model

Markov Logic Networks

$$p(G,L,S,Z,W) = \frac{1}{Z}\exp(\beta_1\phi_1(G,L,S,Z,W) + \beta_2\phi_2(G,L,S,Z,W) + \ \ldots)$$

Probability of a model

weighted satisfaction of logical rules

(e.g. `2.75:: L -> Z XOR S`)

The use of logic: Model

Markov Logic Networks

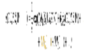
$$p(G,L,S,Z,W) = \frac{1}{Z}\exp(\beta_1\phi_1(G,L,S,Z,W) + \beta_2\phi_2(G,L,S,Z,W) + \ldots)$$

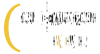
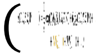
The use of logic: Model

Relational Neural Machines

Add neural-unary factors to MLN

conditioning on subsymbols

$p(G, L, S, Z, W \mid$  $) = \frac{1}{Z} \exp(\beta_1 \phi_1(G, L, S, Z, W) + \beta_2 \phi_2(G, L, S, Z, W) + \ldots$

$+ \; \boldsymbol{NN_G}($  $) + NN_L($  $) + \ldots)$

neural unary factors

Marra et al, ECML 2019
Marra et al., ECAI 2020

The use of logic: Model

Relational Neural Machines

conditioning on subsymbols

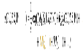
$$p(G, L, S, Z, W \mid \quad) = \frac{1}{Z} \exp(\beta_1 \phi_1(G, L, S, Z, W) + \beta_2 \phi_2(G, L, S, Z, W) + \ldots$$
$$+ \; NN_G(\quad) + NN_L(\quad) + \; \ldots \;)$$

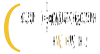
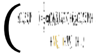
neural unary factors

- Can deal with perception
- Approximate inference is helped by NN

The use of logic: Model

Relational Neural Machines

$p(G, L, S, Z, W \mid$  $) = \frac{1}{Z} \exp(\beta_1 \phi_1(G, L, S, Z, W) + \beta_2 \phi_2(G, L, S, Z, W) + \ldots$

$+ \boldsymbol{NN_G}($  $) + NN_L($  $) + \ldots)$

Neural only unary!

The use of logic: Model
Neural Markov Logic Networks

Add neural relational factors to MLN

$$p(G, L, S, Z, W) = \frac{1}{Z}\exp(\beta_1 \boldsymbol{NN_1(G, L, S, Z, W)} + \beta_2 NN_2(G, L, S, Z, W) + \ \ldots)$$

neural factors also among symbols

Marra et al, UAI 2021
Jung et al, IJAR, 2024

The use of logic: Model

Neural Markov Logic Networks

Can deal with partial or no knowledge (related to structure learning)

$$p(G, L, S, Z, W) = \frac{1}{Z} \exp(\beta_1 NN_1(G, L, S, Z, W) + \beta_2 NN_2(G, L, S, Z, W) + \ldots)$$

neural factors also among symbols

Marra et al, UAI 2021
Svatoš, et al, ILP 2022

The use of logic: Model

Quantified NMLNs

$$p(w) = \frac{1}{Z}\exp(\beta_1 \sum_{\gamma\in\omega} NN_1(\gamma) + \beta_2 \sum_{\gamma\in\omega} NN_2(\gamma) + \dots)$$

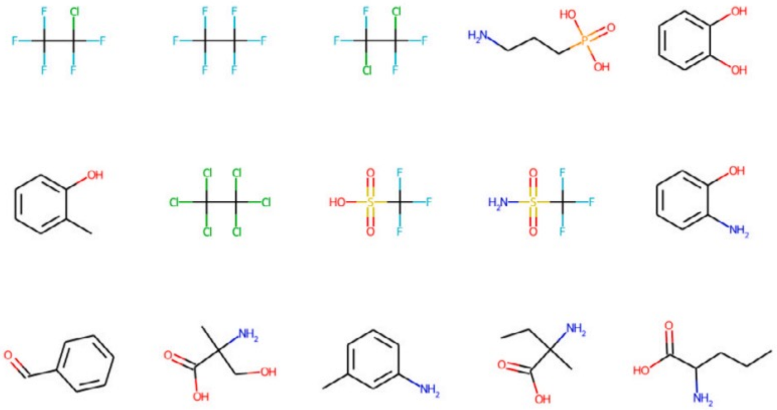
$$p(w) = \frac{1}{Z}\exp(\beta_1 \boldsymbol{A_\gamma} NN_1(\gamma) + \beta_2 \boldsymbol{A_\gamma} NN_2(\gamma) \dots)$$

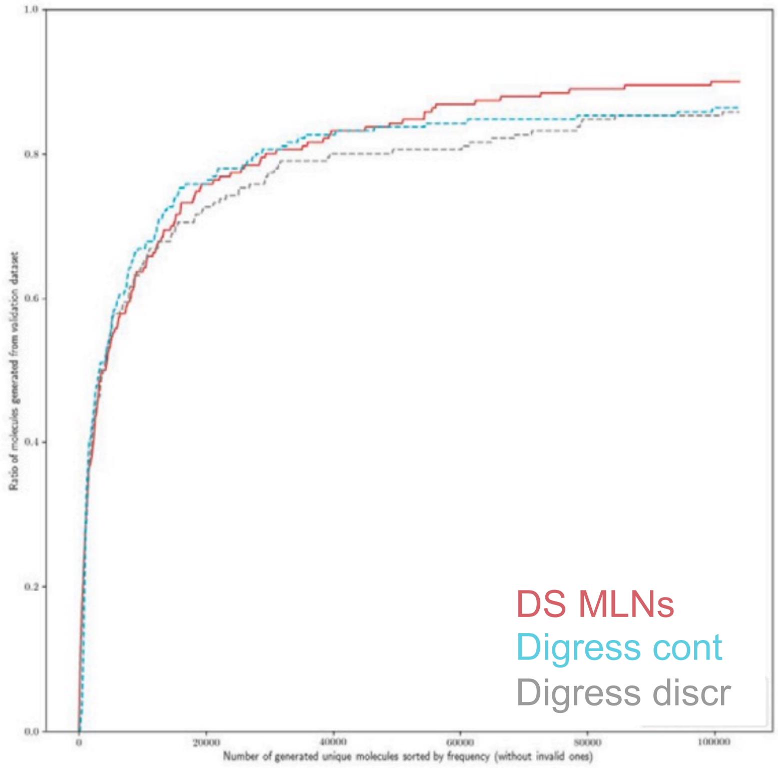
$$p(w) = \frac{1}{Z}\exp(\beta_1 \boldsymbol{DS}\big(NN_1(\gamma)\big) + \beta_2 \boldsymbol{DS}(NN_2(\gamma)) \dots)$$

Jung et al, IJAR 2024

The use of logic: Model

Quantified NMLNs





Jung et al, IJAR 2024

The use of logic: Model
MinMax Entropy Models

All the previous models are solutions of the same optimization problem

min **max** Entropy of $p(G, L, S, Z, W \mid$  $)$
s.t.

- Satisfaction of Logical Constraints w.r.t. data
- Satisfaction **Neural** Constraints w.r.t data
  - unary (RNM)
  - relational (NMLN)

Marra, 2020, PhD Thesis

The use of logic: Proof

The use of logic: Proof

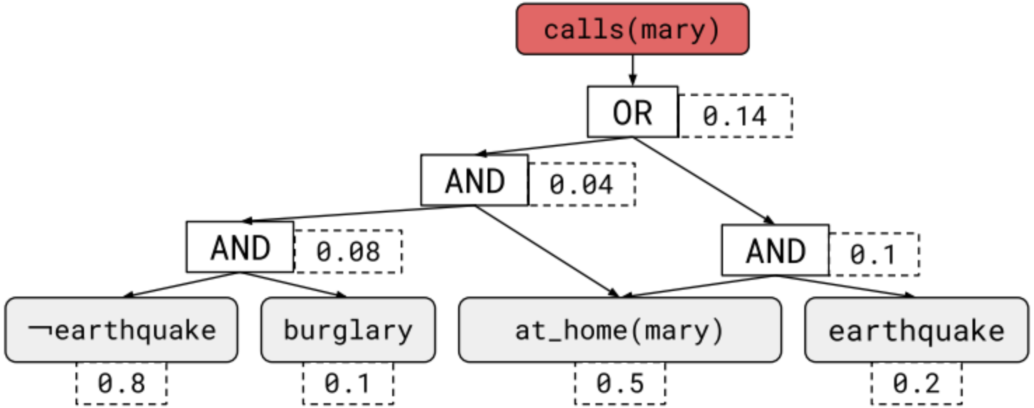
ProbLog

```
0.2::earthquake.
0.1::burglary.
0.5::at_home(mary).
0.4::at_home(john).
alarm :- earthquake.
alarm :- burglary.
calls(X):-alarm,at_home(X).
```

The use of logic: Proof

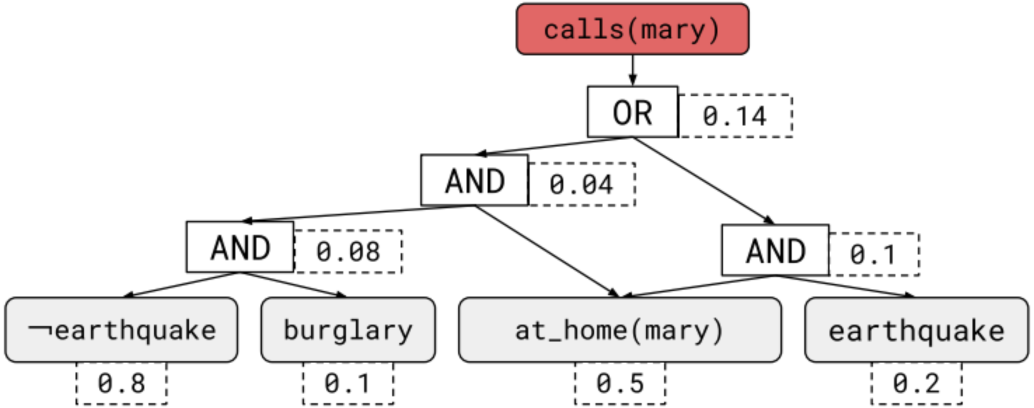
ProbLog

```
0.2::earthquake.
0.1::burglary.
0.5::at_home(mary).
0.4::at_home(john).
alarm :- earthquake.
alarm :- burglary.
calls(X):-alarm,at_home(X).
```



The use of logic: Proof
ProbLog

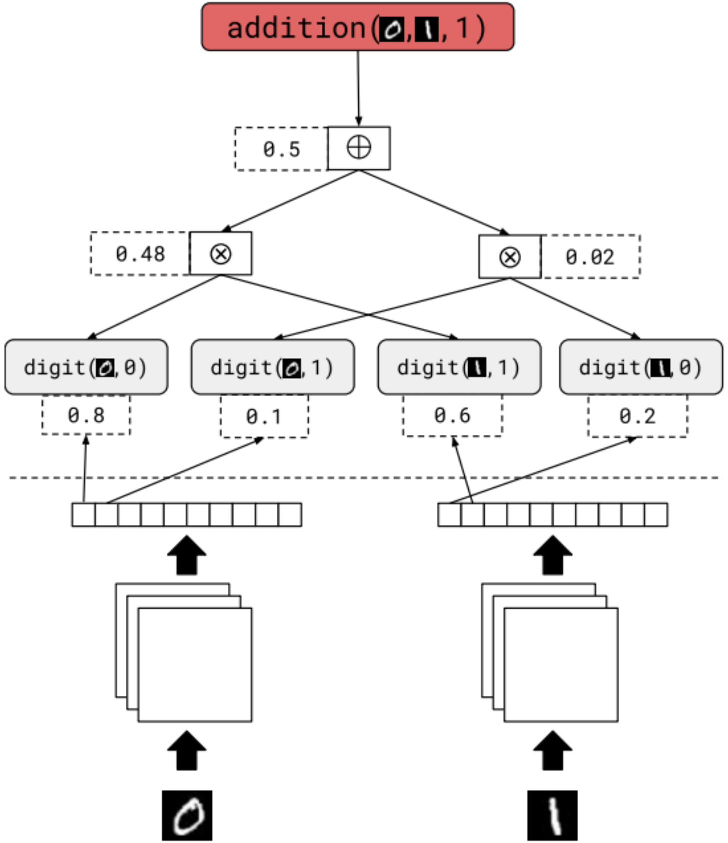
```
0.2::earthquake.
0.1::burglary.
0.5::at_home(mary).
0.4::at_home(john).
alarm :- earthquake.
alarm :- burglary.
calls(X):-alarm,at_home(X).
```



$$p(calls(mary)) = \sum_{\omega} p(earthquake_{\omega}) \cdot p(burglary_{\omega}) \ldots$$

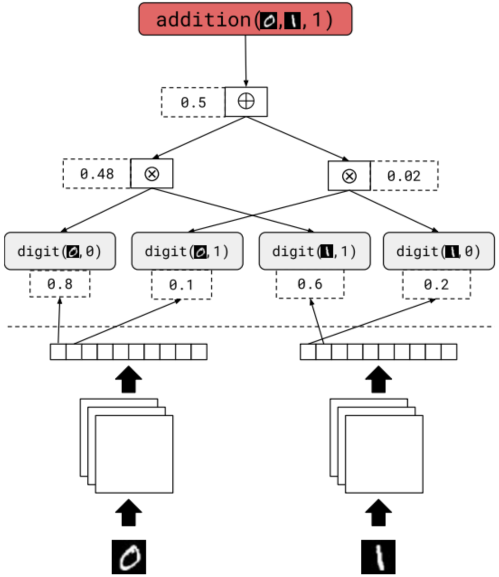
The use of logic: Proof
DeepProbLog

```
nn(m_digit, [X], Y, [0...9]) :: digit(X,Y).
addition(X,Y,Z) :- digit(X,N1), digit(Y,N2), Z is N1+N2.
```



Manhaeve et al, NeurIPS 2018, AIJ 2021, KR 2022

The use of logic: Proof
DeepProbLog




```
nn(m_digit, [X], Y, [0...9]) :: digit(X,Y).
addition(X,Y,Z) :- digit(X,N1), digit(Y,N2), Z is N1+N2.
```

$$p(addition(\; , \; ,1)) = \sum_{\omega} p(digit_{\omega}(X,0) | \;) \cdot p(digit_{\omega}(X,1) | \;) \ldots$$

Neural Predicates

Manhaeve et al, NeurIPS 2018, AIJ 2021, KR 2022

The use of logic: Proof

Stochastic Definite Clause Grammars

Parse Sequences: [“0”, “+”, “9”, “+”, “1”].

```
0.5 :: e(N) --> n(N).
0.5 :: e(N) --> e(N1), p, n(N2),
                {N is N1 + N2}.
1.0 :: p    --> [“+”].

0.1 :: n(0) --> [“0”].
0.1 :: n(1) --> [“1”].
       …
0.1 :: n(9) --> [“9”].
```

The use of logic: Proof

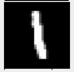
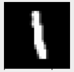
DeepStochLog

0.5 :: e(N) --> n(N).

0.5 :: e(N) --> e(N1), p, n(N2),
{N is N1 + N2}.

nn(,“+”):: p --> [].

nn(,0):: n(0) --> [].

nn(,1):: n(1) --> [].

...

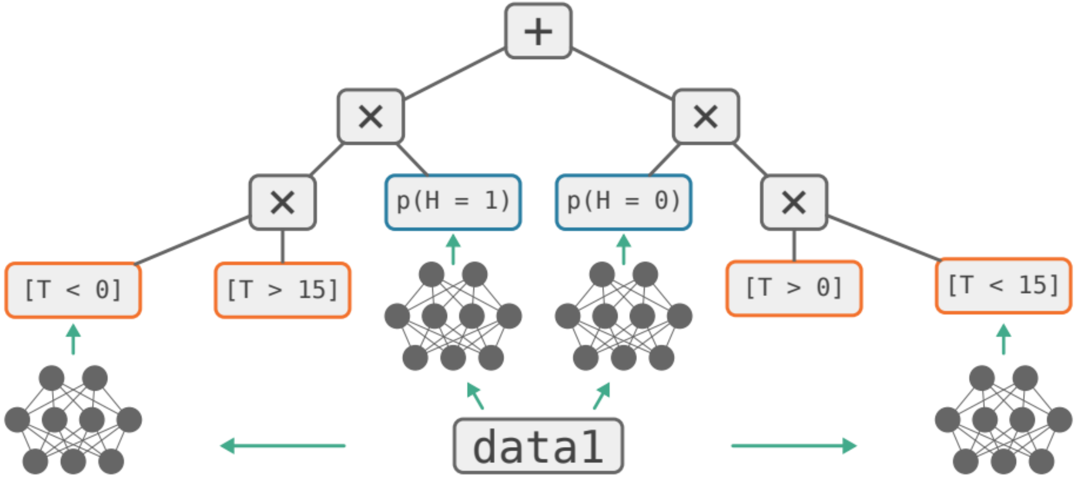
nn(,9):: n(9) --> [].

neural rule

Winters,Marra,
Manhaeve,De Raedt,
AAAI 2022

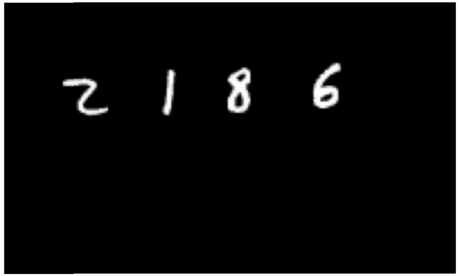
The use of logic: Proof
DeepSeaProbLog

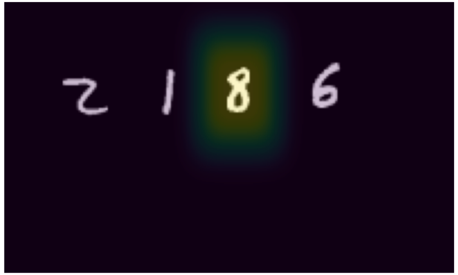
```
humid(Data) ~
    bernoulli(humid_detector(Data)).
temp(Data, T) ~
    normal(temperature_predictor(Data)).

good_weather(Data):-
    humid(Data) =:= 1, temp(Data) < 0.

good_weather(Data):-
    humid(Data) =:= 0, temp(Data) > 15.

query(good_weather(🌎)).
```



De Smet et al, UAI 2023

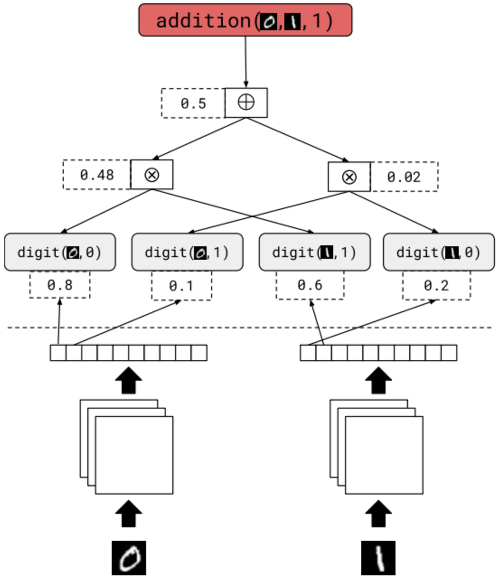
The use of logic: Proof
DeepSeaProbLog





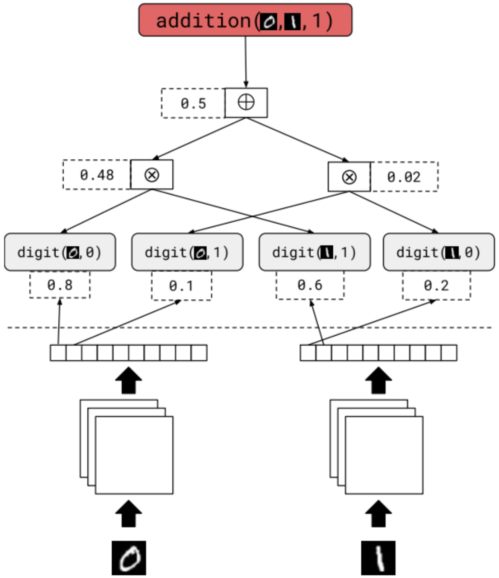
De Smet et al, UAI 2023

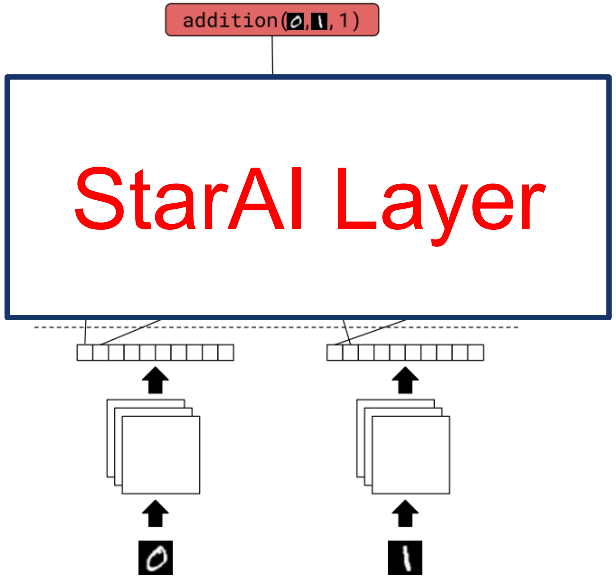
The use of logic: Proof



The use of logic: Proof

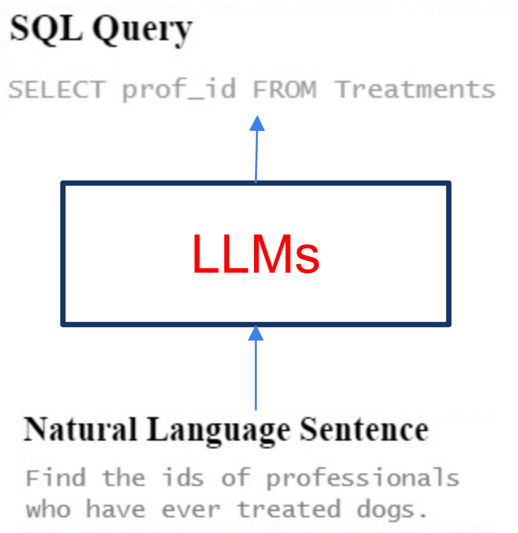
Logic as a layer





The use of logic: Proof

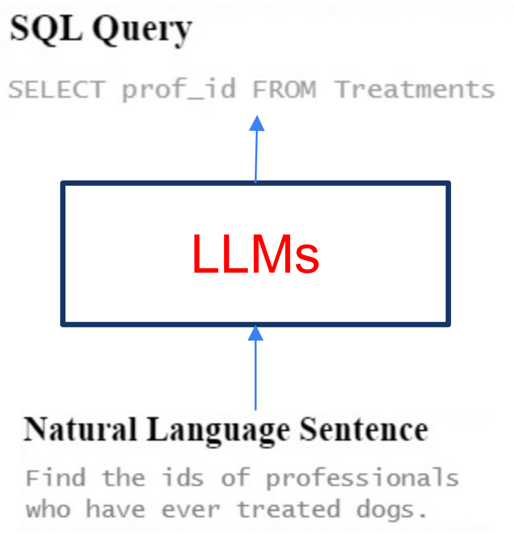
Logic as a layer

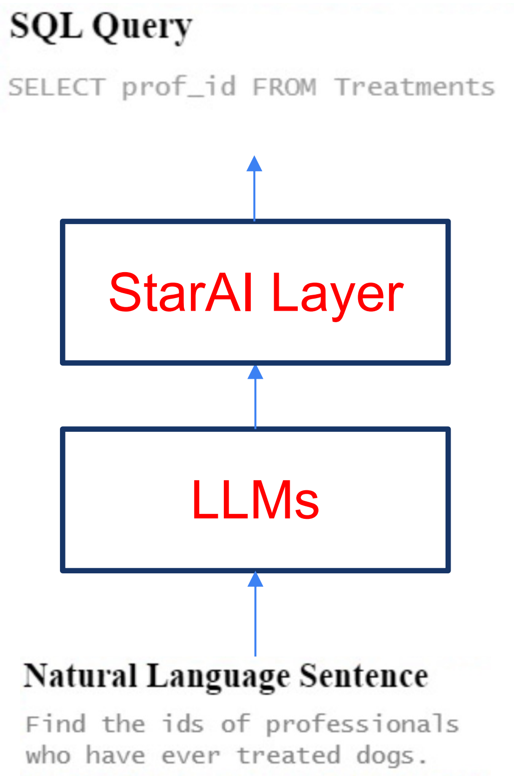


+ (poss.) schema

The use of logic: Proof

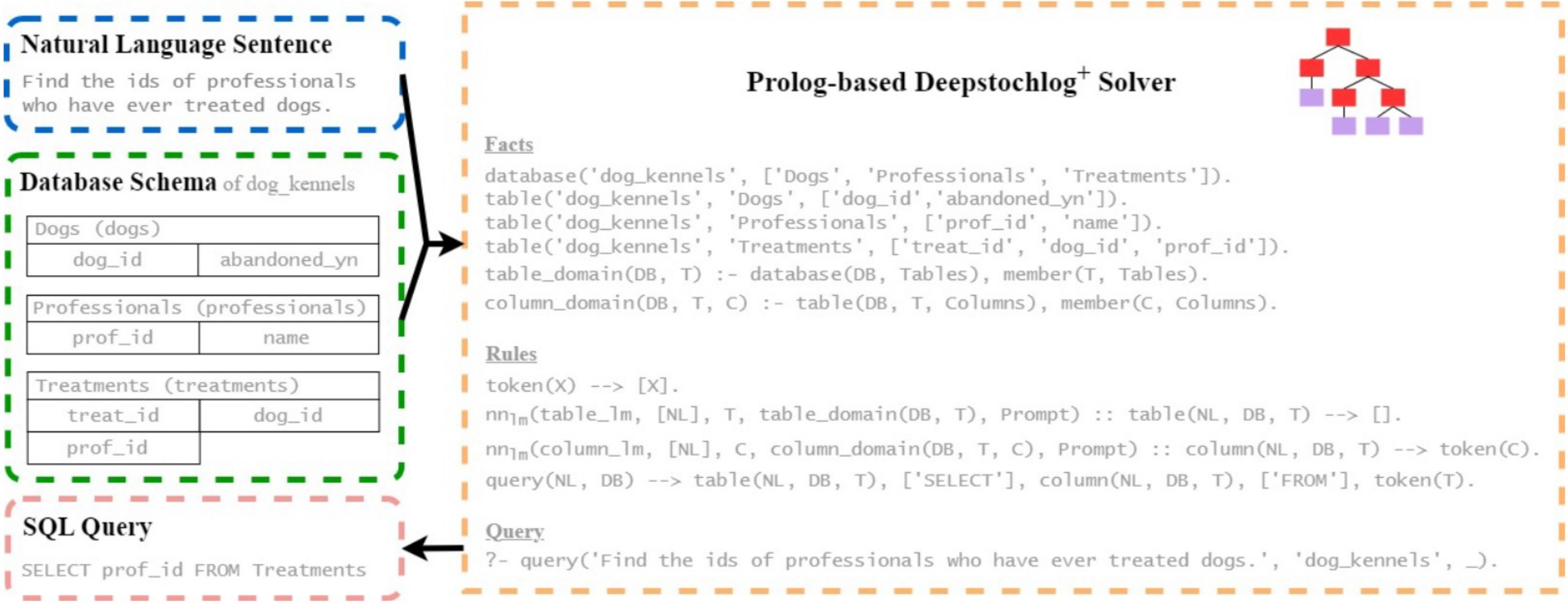
Logic as a layer





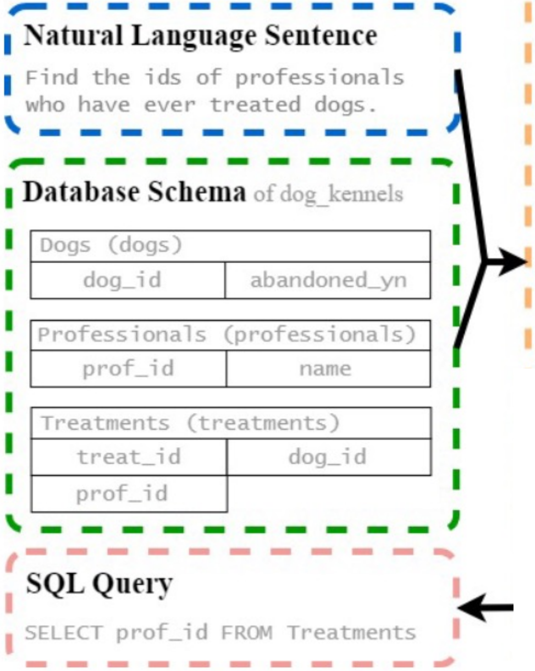
The use of logic: Proof

DeepStochLog



The use of logic: Proof

DeepStochLog



Prolog-based Deepstochlog^{+} Solver

Facts

```
database('dog_kennels', ['Dogs', 'Professionals', 'Treatments']).
table('dog_kennels', 'Dogs', ['dog_id','abandoned_yn']).
table('dog_kennels', 'Professionals', ['prof_id', 'name']).
table('dog_kennels', 'Treatments', ['treat_id', 'dog_id', 'prof_id']).
table_domain(DB, T) :- database(DB, Tables), member(T, Tables).
column_domain(DB, T, C) :- table(DB, T, Columns), member(C, Columns).
```

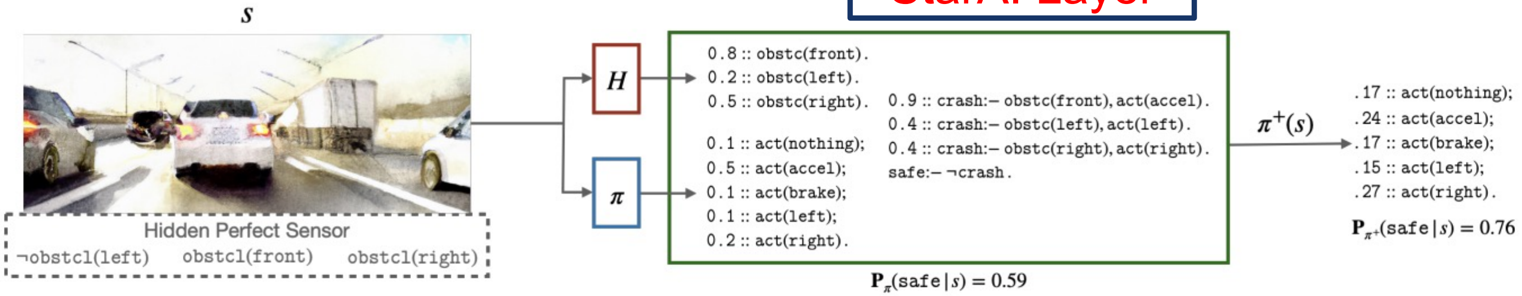
		Validity%	Exact Matching %
Smaller Models (Millions Params.)	T5-small	53.9	41.1
	T5-small+CFGs	88.8	67.1
	Ours (T5-small+DCGs)	**100.0**	75.6
Larger Models (B/Trillions Params.)	DAIL-SQL (GPT-4)	99.2	88.8
	DIN-SQL (GPT-4)	99.2	78.7
	Graphix-T5 (T5-3B+PICARD)	99.6	**91.9**

rule

The use of logic: Proof

Logic layers in policy gradient

StarAI Layer

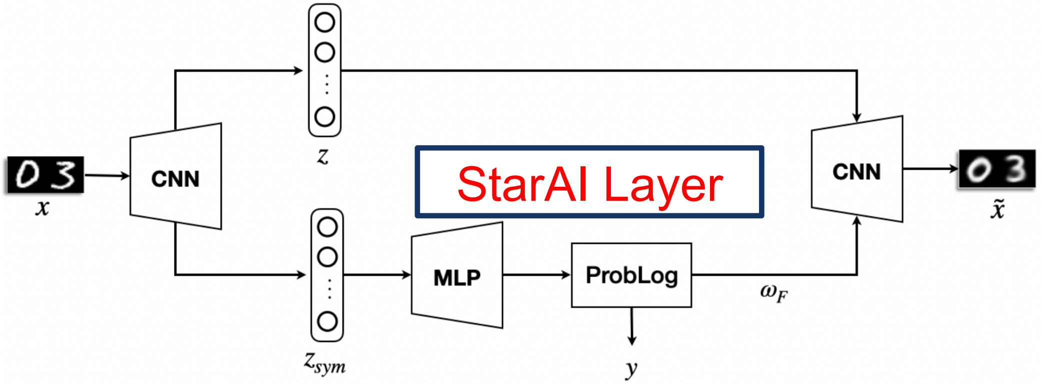


Probabilistic Logic Shields for Safe Reinforcement Learning

Yang et al, IJCAI 2023

The use of logic: Proof

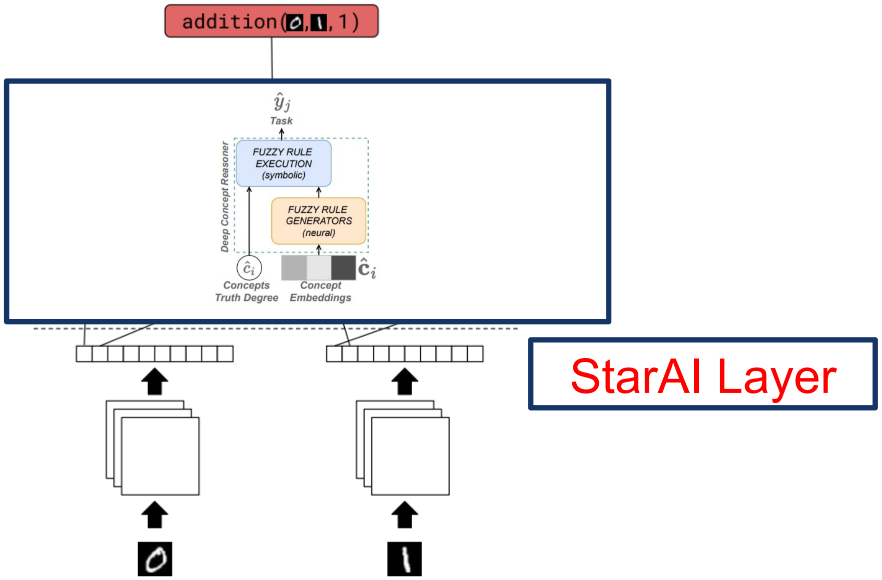
Logic layers to condition VAE



Zero shot generaliation
by programming VAEs

Misino et al, NeurIPS 2022

The use of logic: Proof
Deep Concept Reasoners



We can substitute the logic
with (interpretable) neural networks

Espinoza et al, NeurIPS 2022
Barbiero et al, ICML 2023

StarAI as a recipe for NeSy

- NeSy can bridge the neural and symbolic traditions
- NeSy is still a mess, hard to compare, lack of semantics
- StarAI has already studied sound semantics for learning and reasoning
- StarAI can be used as a starting point for NeSy

StarAI as a recipe for NeSy

StarAI

NeSy

```
p(symbols; params)
```

```
p(symbols; neural nets)
```

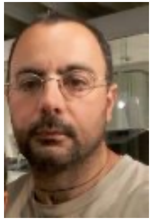
Joint work

DTAI
KU Leuven

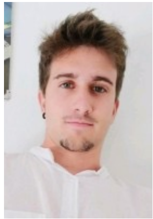
Luc De Raedt
Robin Manhaeve
Thomas Winters
Wen-Chi Yang
Lennert De Smet

SAILAB
Univ. of Siena

Michelangelo Diligenti
Francesco Giannini
Marco Gori
Marco Maggini

elsewhere

Ondrej Kuzelka
Sebastijan Dumancic
Pietro Barbiero
Gabriele Ciravegna
Eleonora Misino
Emanuele Sansone
Gavin Rens

Thank you!