36th Italian Conference on Computational Logic (CILC)

Parma (and virtual), September 8th, 2021

Neuro-Symbolic Learning and Reasoning for Natural Language Processing Tasks

Marco Lippi

marco.lippi@unimore.it

UNIMORE
UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA

Introduction and Motivation

The connectionist vs. symbolic dilemma

A central question in AI

How is knowledge represented in our mind ?

Symbolic approaches

- Reasoning as the result of formal manipulation of symbols

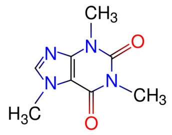
Connectionist (sub-symbolic) approaches

- Reasoning as the result of processing of interconnected (networks of) simple units

Connectionist vs. symbolic approaches

Symbolic approaches

- founded on the **principles of logic**
- exploiting **background knowledge**
- highly **interpretable**

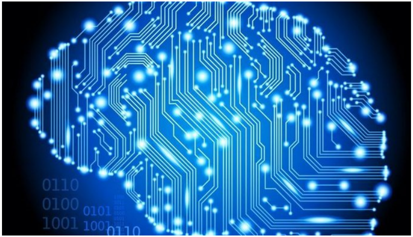


toxic(m) :- doublebond(m,c1,c2), hydroxyl(c2), methyl(m)

Connectionist approaches

- can more easily deal with **uncertain knowledge**
- can be easily **distributed**
- often seen as "black box" → dark magic ☺

Deep learning



Deep learning has brought (back?) a **revolution** into AI

- exploit **more computational power**
- refine **optimization** methods (dropout, rectification, ...)
- automatically **learn feature** hierarchies
- exploit **unsupervised** data (though not yet enough)

Deep learning

Breakthough in a variety of application fields

- Speech recognition
- Computer vision
- Natural language processing
- ...

Is this the **solution** to all AI problems? Probably not but...

- for certain types of task **it is hard to compete**
- **big companies** are currently playing a major role
- huge space for applications **upon deep learning systems**

What is missing?

Deep learning

Still clearly a **sub-symbolic approach**

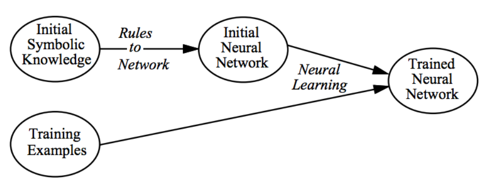
- Building models that are hard to **interpret**
- **Representation learning**: a step towards symbols
- **eXplainable AI**: make deep networks interpretable
- Is there any connection with symbolic approaches?
- What about **logic and reasoning**?

Is it possible to combine **both** worlds?

Pioneering approaches

Knowledge-based artificial neural networks (KBANNs)

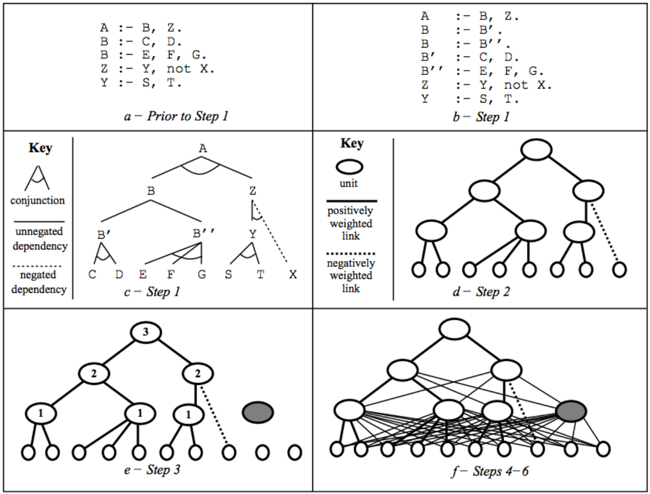
- [Towell & Shavlik, 1994]
- One of the first attempts to inject knowledge into ANNs
- Trying to interpret an ANN model as a set of logic rules



Knowledge Base		*Neural Network*
Final Conclusions	$\Longleftrightarrow$	Output Units
Supporting Facts	$\Longleftrightarrow$	Input Units
Intermediate Conclusions	$\Longleftrightarrow$	Hidden Units
Dependencies	$\Longleftrightarrow$	Weighted Connections

Pioneering approaches

KBANN: an example



NeSy and SRL

More recent research directions:

- **Statistical Relational Learning** (SRL)
- **Neural-Symbolic Learning and Reasoning** (NeSy)

$\rightarrow$ developed during the 90s-00s
$\rightarrow$ combining logic with probabilistic/statistical learning (SRL)
$\rightarrow$ combining logic with cognitive neuroscience (NeSy)

Another recent research direction: **purely sub-symbolic** approaches also to include **background knowledge** and to perform **reasoning tasks**

SRL

Research area that aims at combining **first-order logic** and **graphical models** for learning and reasoning

- Exploit the **expressive power** of first-order logic
- Handle uncertainty with **graphical models**
- Combine **logic and probabilistic** inference

NeSy

Research area that aims at combining **neural models** and **symbolic approaches** for learning and reasoning

- Encode knowledge in the **architecture** of the network
- Use a **regularization** term to encode rules
- **Constrain** neural computations with rules

Caveat [De Raedt, 2020]: inject knowledge into the neural network, then let the network do the rest might not be sufficient $\rightarrow$ partly lost the power of reasoning and explanation

Caveat [Bengio, 2021]: neural computations are necessary to ensure the scalability of both learning and reasoning tasks

Neuro-Symbolic AI

A wide plethora of alternatives

- Directed vs. undirected **graphical models**
- Grounding vs. proofs for **inference**
- **Learning** parameters and/or structure
- Different **types of logic**
- Different uses of **background knowledge**

NeSy: some examples of frameworks

Markov Logic Networks [Richardson & Domingos, 2006]

A probabilistic-logic framework to model knowledge

An example

```
Movie = {BladeRunner, TheMatrix}
Person = {Alice, Bob, Carl, David}
2.3  LikesMovie(x,m) ∧ Friends(x,y) => LikesMovie(y,m)
1.6  Friends(x,y) ∧ Friends(y,z) => Friends(x,z)
```

The **higher** is the weight, the **more likely** is a world where the rule is true, other things being equal.

Markov Logic Networks [Richardson & Domingos, 2006]

The **probability** of a world/configuration depends on the **weights** (w_i) and the number of **groundings** (n_i) of each formula (F_i):

$$P(Y = y | X = x) = \frac{\exp(\sum_{F_i \in \mathcal{F}} w_i n_i(x, y))}{Z_x}$$

Inference aims to find the **most probable** y given x:

$$y^* = \text{argmax}_y P(Y = y | X = x)$$

Markov Logic Networks [Richardson & Domingos, 2006]

Learning

Both **weights** and **rules** themselves can be **learned** from a collection of predicate observations.

Inference

Given a set of known facts, the weighted rules can be used to infer the truth value of other (query) facts.

An example

`LikesMovie(Alice,BladeRunner)`

`Friends(Alice,Bob)`

`Friends(Alice,Carl)`

`LikesMovie(Carl,BladeRunner)???`

Ground-Specific MLNs [Lippi & Frasconi, 2009]

An extension of Markov Logic Networks that allows to embed neural networks to compute weights

A simple classification example

```
w(x)   HasFeatures(x,$f) => PositiveClass(x)
```

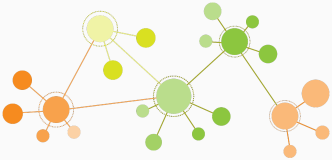
The weight `w(x)` is computed by a neural network using (any) set of features `$f` describing example `x`.

These are named **Ground-Specific MLNs**.

Collective classification

This framework could be easily exploited to perform collective classification on a set of **non-independent examples**, like nodes in a graph, agents in a network, sentences in a document, ...

Given a set of (possibly neural) rules, and a collection of constants/features representing the document, the inference algorithm computes the **most likely world**, or interpretation, thus assigning a truth value to each predicate in the document.



Ground-Specific MLNs [Lippi & Frasconi, 2009]

An example in structured text classification

```
 2.3   Features(X,$F1) => CategoryA(X)
-1.8   Features(X,$F1) => CategoryB(X)
 0.9   Features(Y,$F2) => CategoryA(Y)
-0.7   Features(Y,$F2) => CategoryB(Y)
 1.1   Features(X,$F1) ∧ Features(Y,$F2) => Link(X,Y)
+Inf   Link(X,Y) => CategoryA(X) ∧ CategoryB(Y)
```

Ground-specific weights are computed by neural networks.

Infinite weights correspond to **hard** constraints.

DeepProbLog [Manhaeve et al., 2018]

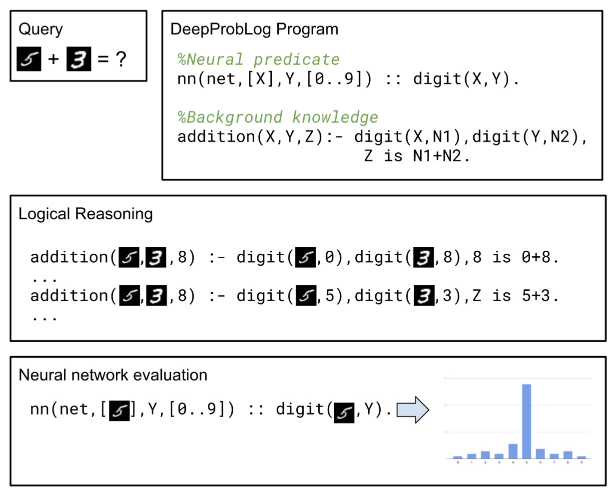
Problog is a **probabilistic extension of Prolog** where probabilities can be attached to ground facts or rules.

DeepProblog extends Problog by computing such probabilities with neural networks, within a framework for probabilistic reasoning

- Necessary to know Pro(b)log
- Cannot (yet) perform collective classification

DeepProbLog [Manhaeve et al., 2018]

An example [Manhaeve et al., 2018]



DeepProbLog [Manhaeve et al., 2018]

An example in structured text classification

```
nn(net1,H,[catA,catB,catC]) ::
  type(H,catA);
  type(H,catB);
  type(H,catC).

nn(net2,H,[link,none]) ::
  type(H,link);
  type(H,none).

type(Y,catA) :- rel(X,Y,link).
type(X,catB) :- rel(X,Y,link).
```

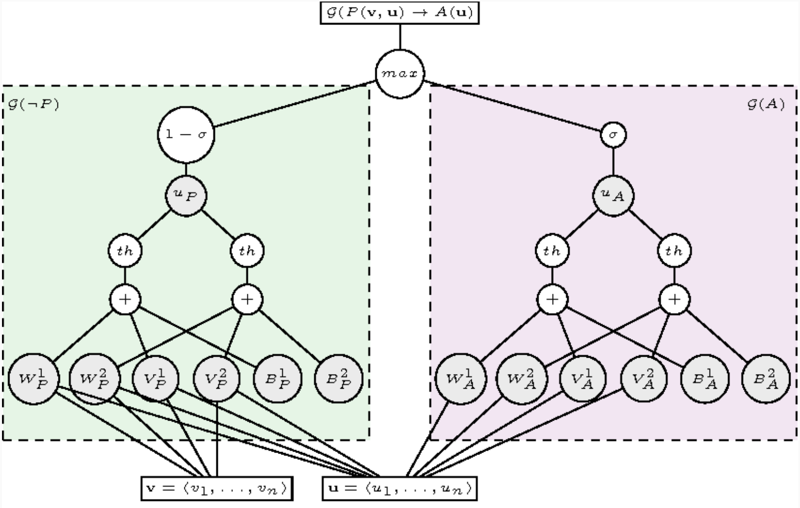
Logic Tensor Networks [D'Avila Garcez & Serafini, 2015]

Framework that combines neural networks with symbolic rules through the use of **fuzzy logic**.

- Use (fuzzy) logic to model background knowledge
- Use deep networks to predict the truth value of predicates
- Translate rules into real-valued functions
- Combine predicates into a tailored network architecture

Logic Tensor Networks [D'Avila Garcez & Serafini, 2015]

An example [Serafini et al., 2016]



Learning from Constraints [Diligenti et al., 2012]

Framework that extends kernel machines as well as neural networks, by including first-order logic clauses in the form of **constraints** within a **regularization term**

- Translate rules into real functions (e.g., p-norms)
- Loss function integrating logic-based penalties
- Penalize solutions where constraints are violated
- Allow collective classification and transductive learning

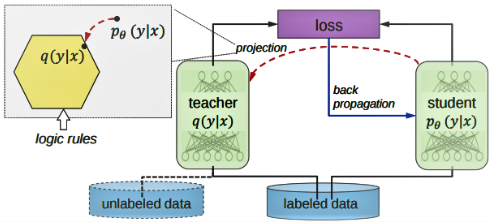
Purely sub-symbolic approaches

Key ideas

- Use some sort of **auxiliary memory**
- **Distillation** from larger to smaller models

Purely sub-symbolic approaches

Harnessing deep neural networks with logic rules



- Rules are encoded in soft logic
- Teacher and student are learned simultaneously
- Knowledge is distilled into the student through the teacher

Applications so far...

- Image classification with rules/taxonomy
- Toy problems for probablistic reasoning
- Link prediction in networks
- Knowledge graph completion
- Semantic Web
- **...What about NLP?**

Applications to NLP (Part I)
Argument Mining

Applications to NLP (Part I)

Argument Mining

Goal: extact arguments from unstructured text

What is an argument? Many models in the literature...

An intuitive definition is given by Douglas Walton:

- a set of **premises**
- a conclusion, sometimes also called **claim**
- an **inference** from the premises to the conclusion

Main tasks for NLP

- Detect argument components
- Detect links between argument components

Applications to NLP (Part I)

An example from the IBM corpus (Wikipedia pages)

CLAIM

Health risks can be produced by long-term use or excessive doses of anabolic steroids

SUPPORTED BY PREMISE

A recent study has also shown that long term anabolic-androgenic steroids (AAS) users were more likely to have symptoms of muscle dysmorphia

Applications to NLP (Part I)

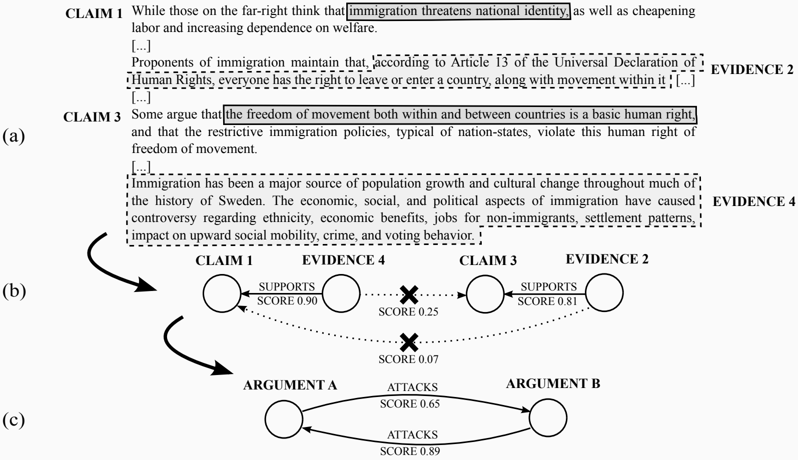


Figure from [Lippi and Torroni, 2016]

Applications to NLP (Part I)

Argument graphs follow **specific rules** that are strictly dependent on the underlying argument model

- If X supports Y, then X is a premise and Y is a claim
- If X supports Y, then Y should not support X
- If X supports Y and Z, then Y should not attack Z
- ...

Other rules can be **soft**, not just hard constraints

- If X and Y are two claims given by two opponent political candidates, it is unlikely that they support each other

Applications to NLP (Part I)

We made some **preliminary experiments** with Logic Tensor Networks (one of their first applications to NLP)

Corpus: **Randomized clinical trials abstracts**

- 659 documents
- three topical datasets: neoplasm, glaucoma, mixed
- 2,808 premises, 1,390 claims
- only 10% of possible pairs are linked

Applications to NLP (Part I)

Argument model used in the corpus:

- links encode **non-symmetric** support relations
- a claim can support only a claim
- an evidence can be supported only by an evidence

LTN rules

$\forall x, y : LINK(x, y) \Rightarrow \sim LINK(y, x)$

$\forall x, y : LINK(x, y) \wedge CLAIM(x) \Rightarrow CLAIM(y)$

$\forall x, y : LINK(x, y) \wedge EVIDENCE(y) \Rightarrow EVIDENCE(x)$

Applications to NLP (Part I)

Dataset	Split	Approach	Classification Comp.	Classification Link	Agreement Comp.	Agreement Link	Properties Eq. 2	Properties Eq. 3
Neoplasm	Validation	Data	83 - 84	42 - 41	77	66	98	**100**
		Data + Rules	84 - **85**	**44** - 43	**81**	**71**	**100**	**100**
Neoplasm	Test	Data	79 - **80**	34 - 31	77	64	98	**100**
		Data + Rules	79 - 78	**35 - 35**	**79**	**70**	**100**	**100**
Glaucoma	Test	Data	**82 - 82**	45 - 43	**75**	66	99	**100**
		Data + Rules	81 - **82**	**47** - 45	**75**	**71**	**100**	**100**
Mixed	Test	Data	**81 - 81**	38 - 34	75	64	98	**100**
		Data + Rules	**81** - 80	39 - **40**	**76**	**69**	**100**	**100**

Table 1: Results of NeSy AM on AbstRCT against the data-driven baseline. For component classification, we report both the result obtained by the MAJ approach (before the dash) and by the AVG approach (after the dash). Scores are reported as percentage values.

Applications to NLP (Part II)
Legal Informatics

Legal Informatics

Detect potentially unfair clauses in online Terms of Service

A task in the direction of **consumer-empowering AI**

Why is a clause potentially unfair for the consumer?

Legal experts exploit their domain knowledge (i.e., the Law) to answer such a question. How to exploit such a knowledge?

Applications to NLP (Part II)

With new products, services, and features launching all the time, we need the flexibility to make changes, impose limits, and occasionally suspend or terminate certain offerings — Endomondo ToS, 2016

Legal rationale: the clause is potentially unfair since the provider has the right for unilateral change of the contract/services/goods/features to maintain a level of flexibility to amend and update services, including discontinuation.

Applications to NLP (Part II)

We focus on 5 unfairness categories

Type of clause	# unfair clauses	% unfair clauses	# rationales
Arbitration (A)	45	0.8	8
Unilateral change (CH)	89	1.7	7
Content removal (CR)	58	1.1	17
Limitation of liability (LTD)	161	3.0	18
Unilateral termination (TER)	121	2.3	28

Applications to NLP (Part II)

Key ideas

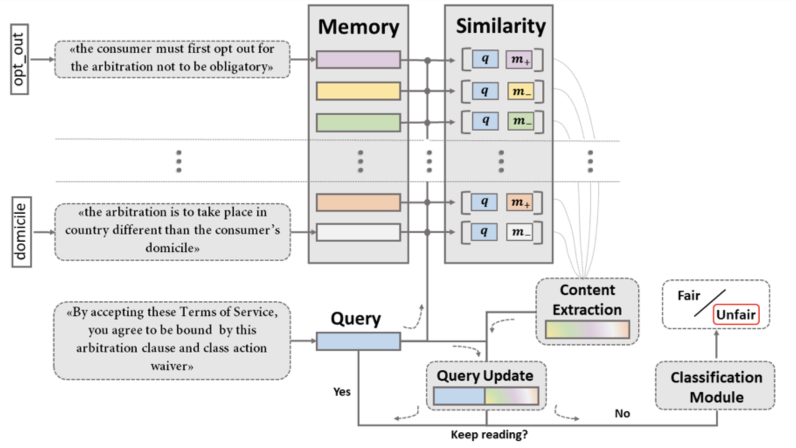
- Use legal **rationales** as background knowledge
- Encode such information into an **external memory**
- Use **attention** to retrieve content from the memory
- Combine **query** and **memory** content to perform classification

Weak vs. strong supervision

- **Weak**: just provide the list of rationales
- **Strong**: link each unfair clause to some rationale

Note: this is a **purely sub-symbolic** approach

A memory-enhanced neural network architecture



Applications to NLP (Part II)

Results on 30 Terms of Service

	A	CH	CR	LTD	TER
No Knowledge					
CNN	0.339	0.506	0.403	0.628	0.583
LSTM	0.302	0.573	0.363	0.602	0.508
DistilBERT	0.447	**0.635**	0.620	0.670	**0.748**
Full Knowledge					
MANN (WS)	0.483	0.506	0.387	0.635	0.602
MANN (SS)	0.465	0.516	0.414	0.605	0.660
MemBERT (WS)	0.494	0.565	0.639	0.664	0.705
MemBERT (SS)	<u>0.504</u>	<u>0.609</u>	**0.670**	**0.686**	<u>0.737</u>
Sampling					
MemBERT (WS) (U-5)	**0.514**	0.556	0.609	<u>0.678</u>	0.702
MemBERT (WS) (P-5-Att-F)	0.491	0.559	0.601	0.643	0.703
MemBERT (WS) (P-5-LG-F)	0.475	0.574	<u>0.660</u>	<u>0.678</u>	0.716
MemBERT (SS) (U-5)	0.503	0.580	0.617	0.652	0.702
MemBERT (SS) (P-5-Att-F)	0.448	0.599	0.635	0.661	0.708
MemBERT (SS) (P-5-LG-F)	0.490	0.536	0.625	0.656	0.706

Results on 30 Terms of Service: interpretability

Table 4: Memory statistics concerning predictions on unfair examples only. Memory interaction is evaluated on ToS-30 test set. We report memory usage (U), the correct memory usage over unfair examples (C) and over examples for which memory is used (CP), along with a more fine-grained ranking version (P@1-3) and the mean reciprocal rank (MRR). Due to different memory usage, columns C–P@3 are not directly comparable. We set activation threshold δ to 0.5.

Model	**U**	**C**	**CP**	**P@1**	**P@3**	**MRR**
Arbitration (A)						
MANN (WS)	0.311	0.289	0.929	0.571	1.000	0.761
MANN (SS)	0.689	0.644	0.935	0.903	0.968	**0.861**
MemBERT (WS)	0.489	0.400	0.818	0.273	0.545	0.478
MemBERT (SS)	0.956	0.911	0.953	0.767	0.837	0.848
Arbitration (CH)						
MANN (WS)	0.169	0.090	0.533	0.000	0.067	0.299
MANN (SS)	0.854	0.730	0.855	0.855	0.961	0.883
MemBERT (WS)	0.404	0.382	0.944	0.250	0.750	0.522
MemBERT (SS)	1.000	0.955	0.955	0.809	0.888	**0.886**

Applications to NLP (Part II)

CLAUDETTE

An Automated Detector of Potentially Unfair Clauses

Claudette found 1 potentially unfair clause (displayed in **bold**) out of 1 sentences.

Hide/show the complete text of the query

Potentially unfair clause #1

We may stop (permanently or temporarily) providing the Services or any features within the Services to you or to users generally .

Unfairness categories: **Unilateral Termination**

Hide/show rationales

The clause is potentially unfair for **Unilateral Termination** since the contract or access can be terminated where the user fails to adhere to its terms, or community standards, or the spirit of the ToS or community terms, including inappropriate behaviour, using cheats or other disallowed practices to improve their situation in the service, deriving disallowed profits from the service, or interfering with other users' enjoyment of the service or otherwise puts them at risk, or is investigated under any suspicion of misconduct. (score = 0.992)

The clause is potentially unfair for **Unilateral Termination** since the contract or access may be terminated where the user has been engaging in illegal or unlawful activity, including fraudulent behaviour, abusive, misusive or otherwise harmful behaviour, or for reasons of safety or fraud prevention (score = 0.770)

The clause is potentially unfair for **Unilateral Termination** since the contract or access may be terminated for any reason, without cause or leaves room for other reasons which are not specified. (score = 0.638)

CLAUDETTE online demo:

`http://claudette.eui.eu/demo`

Conclusions and References

Conclusions

Neuro-symbolic AI is a rapidly evolving area!
Many challenges ahead, still a lot to be done...

- Many approaches without a clear **taxonomy**
- No **off-the-shelf tool** ready for any use
- **Scalability issues** for inference and learning
- Moving towards **eXplainable AI**
- Few benchmarks, few **comparisons** between approaches
- Applications to **real-world problems** and domains

References (1/2)

De Raedt L., Dumancic S., Manhaeve R., Marra G., From Statistical Relational to Neuro-Symbolic Artificial Intelligence, IJCAI 2020

Diligenti M., Gori M., Maggini M., Rigutini L., Bridging logic and kernel machines, Machine Learning 2012

Galassi, A., Lippi, M., Torroni, P., Investigating Logic Tensor Networksfor Neural-Symbolic Argument Mining, submitted 2021

Hu Z., Ma X., Liu Z., Hovy E., Xing E., Harnessing Deep Neural Networks with Logic Rules, ACL 2016

Lippi M., Frasconi P., Prediction of protein β-residue contacts by Markov logic networks with grounding-specific weights, Bioinformatics 2009

References (1/2)

Manhaeve R., Dumancic S., Kimmig A., Demesteer T., De Raedt, L., DeepProbLog: Neural Probabilistic Logic Programming, NeurIPS 2018

Marra G., Giannini F., Diligenti M., Gori M., LYRICS: a General Interface Layer to Integrate Logic Inference and Deep Learning, arXiv 2019

Richardson M., Domingos P., Markov Logic Networks, Machine Learning 2006

Ruggeri F., Lagioia F., Lippi M., Torroni P., Detecting and explaining unfairness in consumer contracts through memory networks, AI&Law 2021

Serafini & D'Avila Garcez, Logic Tensor Networks: Deep Learning and Logical Reasoning from Data and Knowledge, NeSy 2016