Symbolic Transfer Learning through Knowledge Manipulation Methods

Matteo Magnini[*]

[*]Dipartimento di Informatica – Scienza e Ingegneria (DISI)
Alma Mater Studiorum – Università di Bologna
matteo.magnini@unibo.it

Doctoral Consortium at the
20th International Conference on
Principles of Knowledge Representation and Reasoning (KR)
September 5th, 2023, Rhodes (Greece)


Matteo Magnini


Roberta Calegari


Andrea Omicini


Michela Milano


Giovanni Ciatto


Sara Montagna


Enrico Denti


Alessandro Ricci


Mirko Viroli


Giuseppe Pisano


Samuele Burattini


Martina Baiardi


Andrea Agiollo


Andrea Rafanelli


Federico Sabbatini

3rd ELP — 1992

8th ECOOP — 1994

1st AAMAS — 2002

8th DALT School — 2011

5th CompleNet — 2014

18th PRIMA — 2015

28th ECAI — 2025

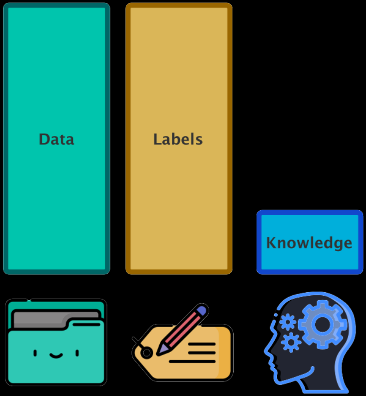
Next in Line. . .

Context

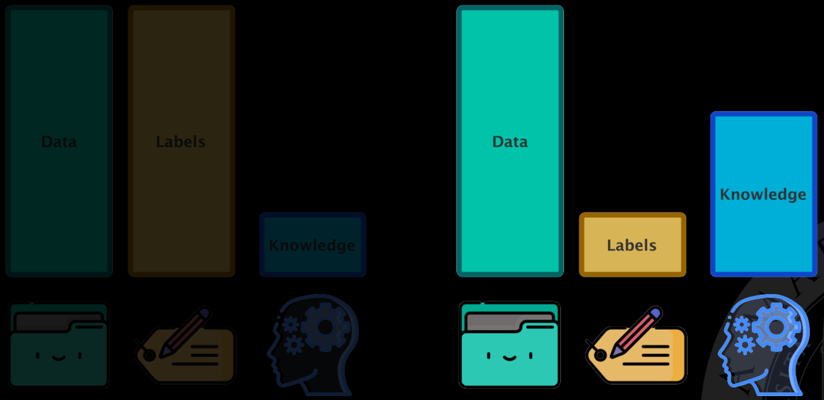
Players involved in a Machine Learning workflow:

- Data
  - ! require labels
- Knowledge
  - human experts
  - common sense
  - extracted from data/models
- Model
  - trained on data
  - sometimes also trained with prior knowledge

Utopy and Reality I



Utopy and Reality II



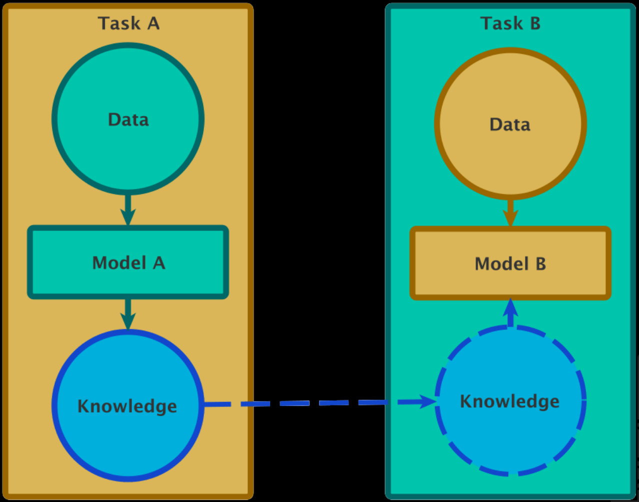
Utopy and Reality III

But in some scenarios we deal with both low data/labels and knowledge. . .

Transfer Learning I

Reuse the knowledge acquired in a similar task! [Tan et al., 2018]

Transfer Learning II



Transfer Learning III

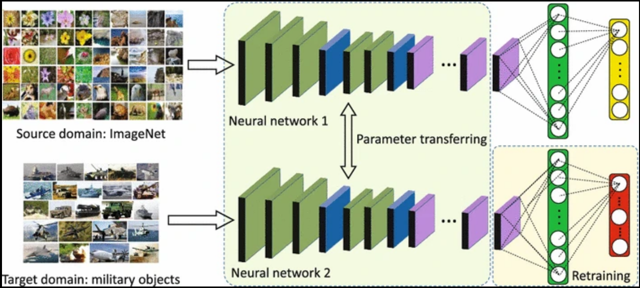
What kind of knowledge?

Transfer Learning IV

What kind of knowledge?

Typically model parameters (NN weights) → sub-symbolic knowledge

Transfer Learning V



source: [Yang et al., 2019]

Next in Line. . .

Limitations of Sub-symbolic TL

- Opaqueness [Brachman and Levesque, 2004]
  - ! sub-symbolic knowledge/model is not interpretable by humans

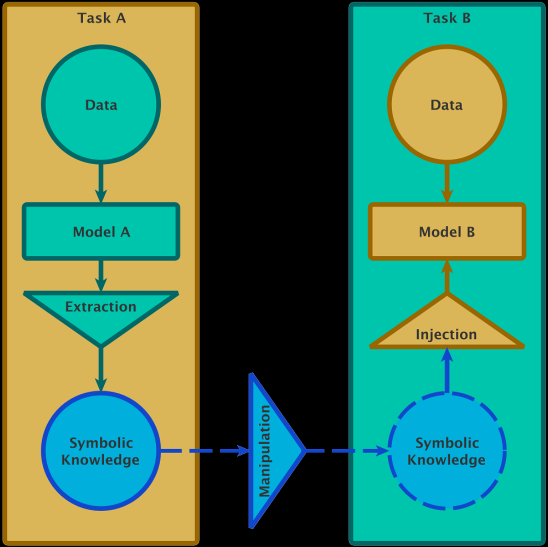
- Why it (does not) works?
  - we have no guarantees on what happens under the hood
  - fails when tasks are different
  - fails when there are biases in the data

Symbolic TL I

A not exhaustive list of benefits of symbolic knowledge: [Besold et al., 2017]

- Interpretability
  - ! symbolic knowledge is understandable by both humans and machines
- Concision
  - intensional representation
  - natural support to recursion
- Lingua Franca
  - no bound to a specific class of models

Symbolic TL II



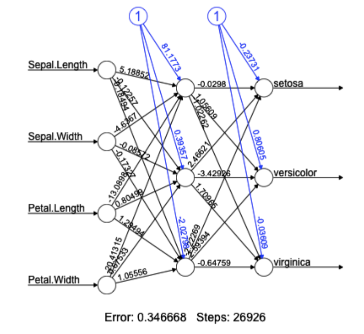
Next in Line...

Symbolic Knowledge Extraction I

Extract symbolic knowledge from a sub-symbolic model: [Guidotti et al., 2019]

- Families
  - pedagogical → model agnostic
  - decompositional → model inspection
- Logical Rules
  - propositional logic
  - first-order logic
  - other logics
- Scope
  - global explanations
  - local explanations

Symbolic Knowledge Extraction II



$\rightarrow$

$$\mathit{Class} = \texttt{setosa} \leftarrow \mathit{PetalWidth} \leq 1.0.$$

$$\mathit{Class} = \texttt{versicolor} \leftarrow \mathit{PetalLength} > 4.9 \wedge \mathit{SepalWidth} \in [2.9,\ 3.2].$$

$$\mathit{Class} = \texttt{versicolor} \leftarrow \mathit{PetalWidth} > 1.6.$$

$$\mathit{Class} = \texttt{virginica} \leftarrow \mathit{SepalWidth} \leq 2.9.$$

$$\mathit{Class} = \texttt{virginica} \leftarrow \mathit{SepalLength} \in [5.4,\ 6.3].$$

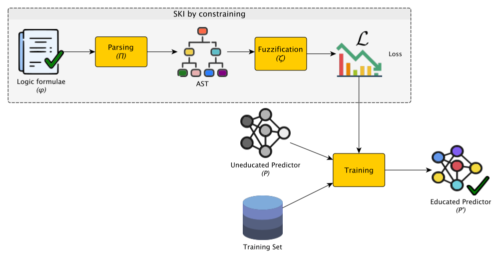
$$\mathit{Class} = \texttt{virginica} \leftarrow \mathit{PetalWidth} \in [1.0,\ 1.6].$$

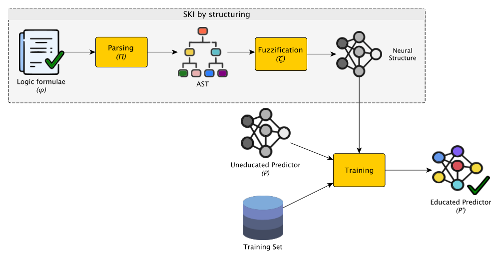
Symbolic Knowledge Injection I

Inject symbolic knowledge into a sub-symbolic model: [von Rueden et al., 2021]

- Families
  - constraining → loss function
  - structuring → model architecture
  - alternatively keep model(s) and knowledge separate
- Logical Rules
  - propositional logic
  - first-order logic
  - other logics

Symbolic Knowledge Injection II





Next in Line...

Open Problems

1. What is in between?
   - it is rare to be able to use the extracted knowledge as is
   - different tasks may require different representations
2. Full recursive concept support?
   - formal logic can represent recursive predicates
   - sub-symbolic models like NN (based on backpropagation) are direct acyclic graphs
3. Evaluation
   - accuracy is not the only dimension
   - need for other metrics (e.g., interpretability, robustness) [Agiollo et al., 2023]
4. Technological support
   - public and maintained tools for symbolic knowledge extraction and injection [Sabbatini et al., 2021, Magnini et al., 2022]

Working Plan

1. Design knowledge manipulation methods:
   - similar tasks
     → partially affects variables and constants
   - dissimilar tasks
     → affects all the knowledge
     → inductive logic programming [Muggleton, 1991]
2. Design injection methods to fully support recursion:
   - new symbolic knowledge injection methods
     → the knowledge is actually injected into the model
   - but also neuro-symbolic approaches
     → model(s) and knowledge are two distinct entities of the system

Symbolic Transfer Learning through Knowledge Manipulation Methods

*Matteo Magnini**

*Dipartimento di Informatica – Scienza e Ingegneria (DISI)
Alma Mater Studiorum – Università di Bologna
matteo.magnini@unibo.it

Doctoral Consortium at the
20th International Conference on
Principles of Knowledge Representation and Reasoning (KR)
September 5th, 2023, Rhodes (Greece)

References I

[Agiollo et al., 2023] Agiollo, A., Rafanelli, A., Magnini, M., Ciatto, G., and Omicini, A. (2023).
Symbolic knowledge injection meets intelligent agents: QoS metrics and experiments.
Autonomous Agents and Multi-Agent Systems, 37(2):27:1–27:30
DOI:10.1007/s10458-023-09609-6

[Besold et al., 2017] Besold, T. R., d'Avila Garcez, A. S., Bader, S., Bowman, H., Domingos, P. M., Hitzler, P., Kühnberger, K., Lamb, L. C., Lowd, D., Lima, P. M. V., de Penning, L., Pinkas, G., Poon, H., and Zaverucha, G. (2017).
Neural-symbolic learning and reasoning: A survey and interpretation.
CoRR, abs/1711.03902
http://arxiv.org/abs/1711.03902

References II

[Brachman and Levesque, 2004] Brachman, R. J. and Levesque, H. J. (2004).
The tradeoff between expressiveness and tractability.
In Brachman, R. J. and Levesque, H. J., editors, *Knowledge Representation and Reasoning*, The Morgan Kaufmann Series in Artificial Intelligence, pages 327–348. Morgan Kaufmann, San Francisco

DOI:https://doi.org/10.1016/B978-155860932-7/50101-1

[Guidotti et al., 2019] Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., and Pedreschi, D. (2019).
A survey of methods for explaining black box models.
ACM Comput. Surv., 51(5):93:1–93:42

DOI:10.1145/3236009

References III

[Magnini et al., 2022] Magnini, M., Ciatto, G., and Omicini, A. (2022).
On the design of PSyKI: a platform for symbolic knowledge injection into sub-symbolic predictors.
In Calvaresi, D., Najjar, A., Winikoff, M., and Främling, K., editors, *Explainable and Transparent AI and Multi-Agent Systems*, volume 13283 of *Lecture Notes in Computer Science*, chapter 6, pages 90–108. Springer
DOI:10.1007/978-3-031-15565-9_6

[Muggleton, 1991] Muggleton, S. H. (1991).
Inductive logic programming.
New Gener. Comput., 8(4):295–318
DOI:10.1007/BF03037089

References IV

[Sabbatini et al., 2021] Sabbatini, F., Ciatto, G., Calegari, R., and Omicini, A. (2021).
On the design of PSyKE: A platform for symbolic knowledge extraction.
In Calegari, R., Ciatto, G., Denti, E., Omicini, A., and Sartor, G., editors, *WOA 2021 – 22nd Workshop "From Objects to Agents"*, volume 2963 of *CEUR Workshop Proceedings*, pages 29–48. Sun SITE Central Europe, RWTH Aachen University
http://ceur-ws.org/Vol-2963/paper14.pdf

[Tan et al., 2018] Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., and Liu, C. (2018).
A survey on deep transfer learning.
In Kůrková, V., Manolopoulos, Y., Hammer, B., Iliadis, L., and Maglogiannis, I., editors, *Artificial Neural Networks and Machine Learning – ICANN 2018*, pages 270–279, Cham. Springer International Publishing
DOI:10.1007/978-3-030-01424-7_27

References V

[von Rueden et al., 2021] von Rueden, L., Mayer, S., Beckh, K., Georgiev, B., Giesselbach, S., Heese, R., Kirsch, B., Walczak, M., Pfrommer, J., Pick, A., Ramamurthy, R., Garcke, J., Bauckhage, C., and Schuecker, J. (2021).
Informed machine learning – a taxonomy and survey of integrating prior knowledge into learning systems.
IEEE Transactions on Knowledge and Data Engineering, pages 1–1
DOI:10.1109/TKDE.2021.3079836

[Yang et al., 2019] Yang, Z., Yu, W., Liang, P., Guo, H., Xia, L., Zhang, F., Ma, Y., and Ma, J. (2019).
Deep transfer learning for military object recognition under small training set condition.
Neural Comput. Appl., 31(10):6469–6478
DOI:10.1007/s00521-018-3468-3