Symbolic Vs Sub-symbolic AI Methods: Friends or Enemies?

Eleni Ilkou[1,2], Maria Koutraki[1,2]

[1]L3S Research Center

[2]Leibniz University of Hannover

CSSA, October 20, 2020

Timeline

- AI periods characterised as summers and winters depending on
  - funding
  - research development
  - technological advancements

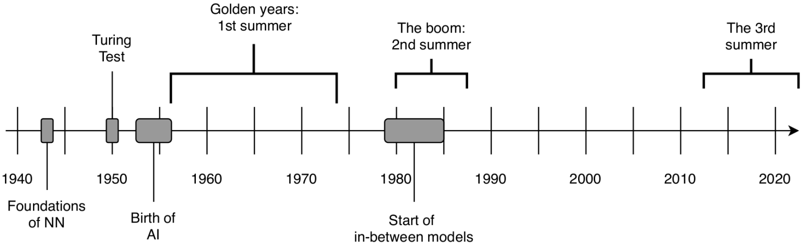


Figure: The timeline of Artificial Intelligence methods

Symbolic AI

Symbolic AI

Advantages:

- explain and reason the results
- human-understandable computation flow
- rule modularity
- inter-operability
- not highly dependent on the input data

Disadvantages:

- datasets with data-quality issues
- prone to noise
- “brittleness”
- high cost of human involvement
- rule bases complex verify and validate

Sub-symbolic AI

Sub-symbolic AI

Advantages:

- robust against noisy and missing data
- high computing performance
- well suitable for big datasets and large KGs
- require less knowledge upfront

Disadvantages:

- not interpretable conclusions
- require huge computation power and huge amounts of data
- biased outcomes

The debate

Symbolic	Sub-symbolic
Symbols	Numbers
Logical	Associative
Serial	Parallel
Reasoning	Learning
von Neumann machines	Dynamic Systems
Localised	Distributed
Rigid and static	Flexible and adaptive
Concept composition and expansion	Concept creation, and generalization
Model abstraction	Fitting to data
Human intervention	Learning from data
Small data	Big data
Literal/precise input	Noisy/incomplete input

Figure: Based on [1, 2] and our analysis

Long and unresolved debate.
The future

- in-between methods

In-between methods

In-between methods

- Discussions started since 1980s while currently there is a high interest in the combination of the fields

- Algorithms based on or have their core on :
  - Neural Network
  - Tensor and Graph Networks
  - Expert Systems
- Characteristics:
  - Algorithms developed for specific applications
  - No need for a-priori assumptions
  - Perform well with noisy data
  - Well fit for large amounts of heterogeneous data

In-between methods in literature

No standard:

- categorisation and taxonomy
- naming method

Techniques in literature:

- connectionist expert systems (or neural network based expert systems)
- multi-agent systems
- hybrid representations
- neural-fuzzy
- neural-symbolic (or neurosymbolic)
  - neurules

Categorisations in Literature

Most categorisations analyse the neurosymbolic range, combination of NN and symbolic methods. We selectively present the categorisations from the works of [3, 4, 5, 6, 7]

CONNECTIONISM	NEUROSYMBOLIC INTEGRATION				SYMBOLICISM
	Unified approaches		Hybrid approaches		
	Neuronal Symbol Proc.	Connectionist Symbol Proc.	Functional hybrids	Translational hybrids	
Segregation	*Neuronal eliminativism*	*Connectionist eliminativism Limitivism Revisionism*	*Hybridization or cohabitation*		*Segregation Implementation−alism*

Figure: The range from symbolic to sub-symbolic as proposed by Hilario [6]

Knowledge Graph applications

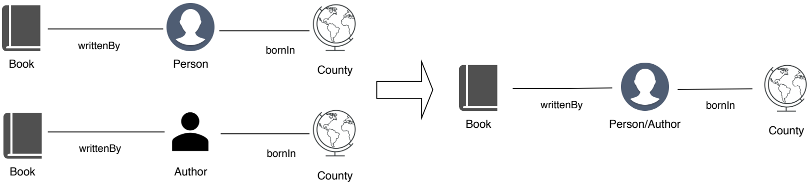
Schema Representation

Traditionally symbolic task, mostly rely on rule mining

- First-order logic
- Ontologies
- Formal knowledge representation languages
  - RDF(S)
  - OWL
  - XML
  - rules

Schema Matching

- Each model uses a usually symbolic input schema
- The majority focuses on class alignment, however, there also are works focus on relation alignment



Schema Matching

Different models to process (matchers):

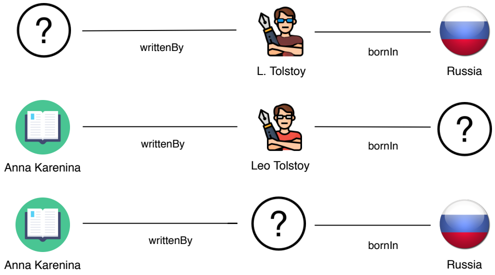
- linguistic or language based (sub-symbolic)
  - combination with NLP
- constrain-based (symbolic)
  - constrains in data features (data types and ranges)
- structured-based (symbolic)
  - focused on database/graph structure

Knowledge Graph Completion (KGC)

Tries to address the:

- missing edges and nodes
- duplicated nodes

Mostly KGEs techniques are used



Entity resolution

Fundamental problem related to data integration

- In 1960s statistical sub-symbolic models
- In 1990s mostly in-between methods
- The techniques usually rely on attribute similarity between the entities
- The algorithms are inspired by IR and relational duplicate elimination

Link prediction techniques

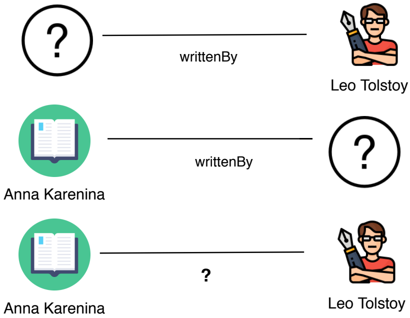


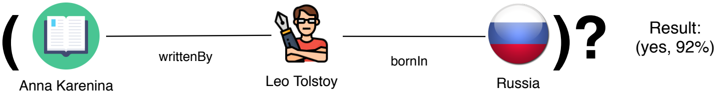
Figure: Head, tail, and relation prediction

Most of SOTA is focused on in-between methods (KGEs, Trans*, neural based KGE with logical rules, and hierarchy-aware KGEs)
Survey of link prediction in complex networks [8]:

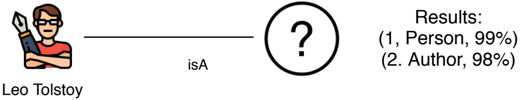
- Only a few are sub-symbolic (ANN, probabilistic and Monte Carlo)
- Most are in-between range

Link prediction tasks

- Triple classification
  - KGEs systems
  - usually in-between methods
  - some use neural tensor networks and time-aware, latent factor and semantic matching models.



- Entity classification
  - related to schema and ontology KG
  - symbolic based [9, 10]



Thank you & contact

For questions or comments please contact on ilkou@l3s.de

Thank you for your attention

Bibliography I

[1] Larry R Medsker. *Hybrid neural network and expert systems*. Springer Science & Business Media, 2012.

[2] Tarek R Besold and Kai-Uwe Kühnberger. Towards integrated neural–symbolic systems for human-level ai: Two research programs helping to bridge the gaps. *Biologically Inspired Cognitive Architectures*, 2015.

[3] Artur d'Avila Garcez, Marco Gori, Luis C Lamb, Luciano Serafini, Michael Spranger, and Son N Tran. Neural-symbolic computing: An effective methodology for principled integration of machine learning and reasoning. *arXiv preprint arXiv:1905.06088*, 2019.

[4] Sebastian Bader and Pascal Hitzler. Dimensions of neural-symbolic integration-a structured survey. *arXiv preprint cs/0511042*, 2005.

[5] Ron Sun and Frederic Alexandre. *Connectionist-symbolic integration: From unified to hybrid approaches*. Psychology Press, 2013.

Bibliography II

[6] Melanie Hilario. An overview of strategies for neurosymbolic integration. *Connectionist-Symbolic Integration: From Unified to Hybrid Approaches*, 1997.

[7] Kenneth McGarry, Stefan Wermter, and John MacIntyre. Hybrid neural systems: from simple coupling to fully integrated neural networks. *Neural Computing Surveys*, 1999.

[8] Babita Pandey, Praveen Kumar Bhanodia, Aditya Khamparia, and Devendra Kumar Pandey. A comprehensive survey of edge prediction in social networks: Techniques, parameters and challenges. *Expert Systems with Applications*, 2019.

[9] Heiko Paulheim and Christian Bizer. Type inference on noisy rdf data. In *International semantic web conference*, 2013.

[10] Jennifer Sleeman and Tim Finin. Type prediction for efficient coreference resolution in heterogeneous semantic graphs. In *2013 IEEE Seventh International Conference on Semantic Computing*, 2013.