Neuro-Symbolic AI for conflict-aware learning over Knowledge Graphs

Laura Balbi

LASIGE Research Unit, Department of Informatics
Faculty of Sciences of the University of Lisbon, Portugal

Contradictions in Knowledge Graphs

Real-world KGs aggregate information from heterogeneous sources (multiple databases, annotators, automated extractors)

This leads to conflicting statements and ambiguous entities/events

! **KG Representation Learning (KGRL) is not able to handle contradictions**

Solution: ignoring contradictions, majority-vote resolution

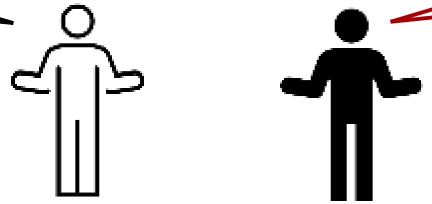
Real-world knowledge inherently includes ambiguity and contradictions

Contradictions aren't just errors – they can ***reflect multiple views*** – valid in their own context/under certain assumptions

e.g. Wikidata [1] entry for **Crimea** (id Q7327):

< Crimea, wiki:country, **Ukraine** >
– start time: 1954

< Crimea, wiki:country, **Russia** >
– start time: 2014



[1] https://www.wikidata.org/

Contradictions in Knowledge Graphs

KGs with explicit contradictions mirror real-world knowledge

But this would mean tackling…

Common assumptions

- **Contradiction → 1 side is wrong**
  (but often both statements are contextually true)
- **Proximity = similarity/relatedness**
  (conflicting edges *close* in the graph will be modeled similarly)

Use-cases with contradictions

Scientific domain – continuously evolving scientific knowledge; different experimental conditions or contexts produce conflicting outcomes

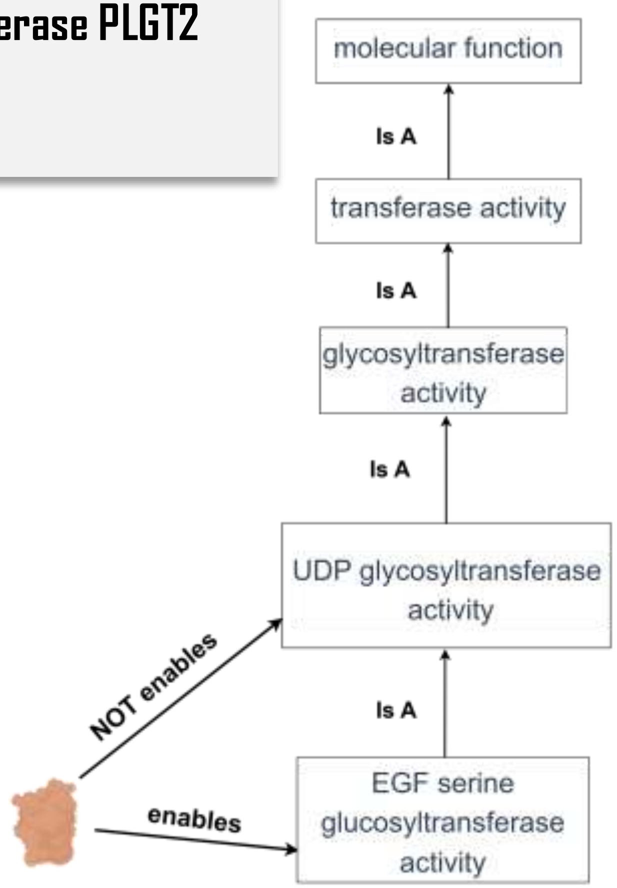
Gene Ontology description for **Transferase PLGT2** (UniProt Q6UW63)

Explicit contradiction

Logical violation of the Gene Ontology's semantics through negation.

Common solution

Detected by reasoners & rule-based logic; valid under specific bio-contexts



General domain – statements from diverse and conflicting viewpoints

Crimea, wiki:country, **Ukraine** — "start time: 1954"

Crimea, wiki:country, **Russia** — "start time: 2014"

Implicit contradiction

Ambiguous value but **no logical violation**: Wikidata's semantics allow for "Crimea" to be associated to more than one country
(no cardinality for wiki:country)

Possible solution

Combining symbolic and sub-symbolic systems to leverage external knowledge as the missing semantics – e.g. LLMs with domain KGs, ontologies.

Hypothesis & Research Questions

Recap

- Contradictions may represent different facets of a complex, multifaceted truth.
- Current KGRL overlooks this and how modeling contradictions explicitly could enhance the applicability of KGs and the accuracy and of ML approaches over KGs.

Hypothesis

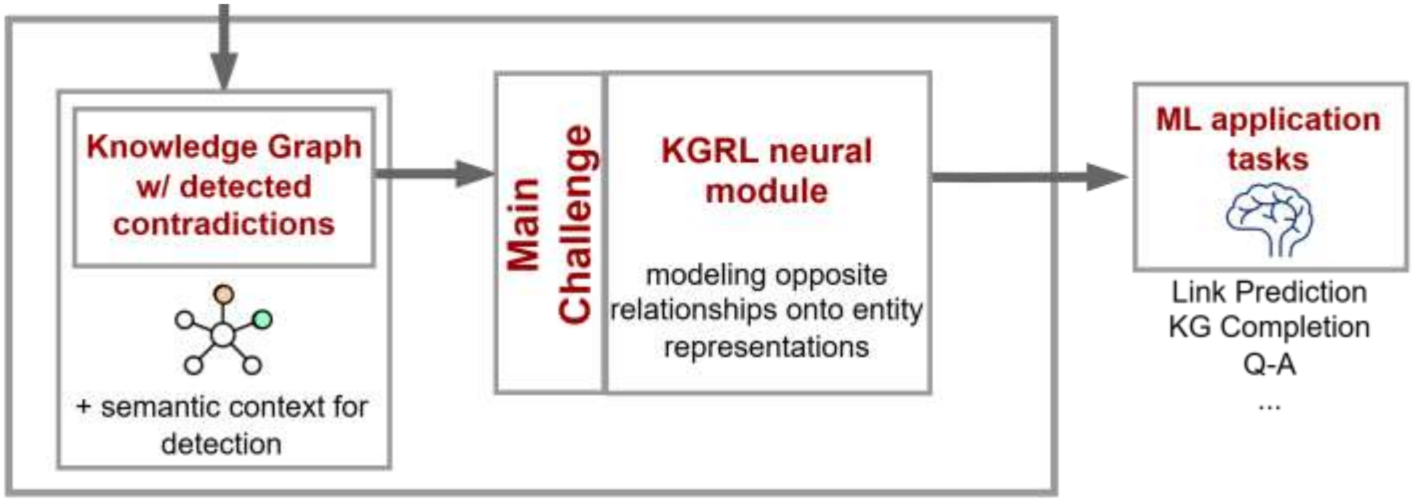
A neuro-symbolic approach that integrates symbolic and sub-symbolic representations can bridge KGs and LLMs to model contradictions and explore them into KG entity representations for use in downstream ML tasks and answer the following research questions:

RQ 1 Does the existence of (different) contradictions in a KG impact the performance, robustness, reliability and transparency of SOTA KGRL?

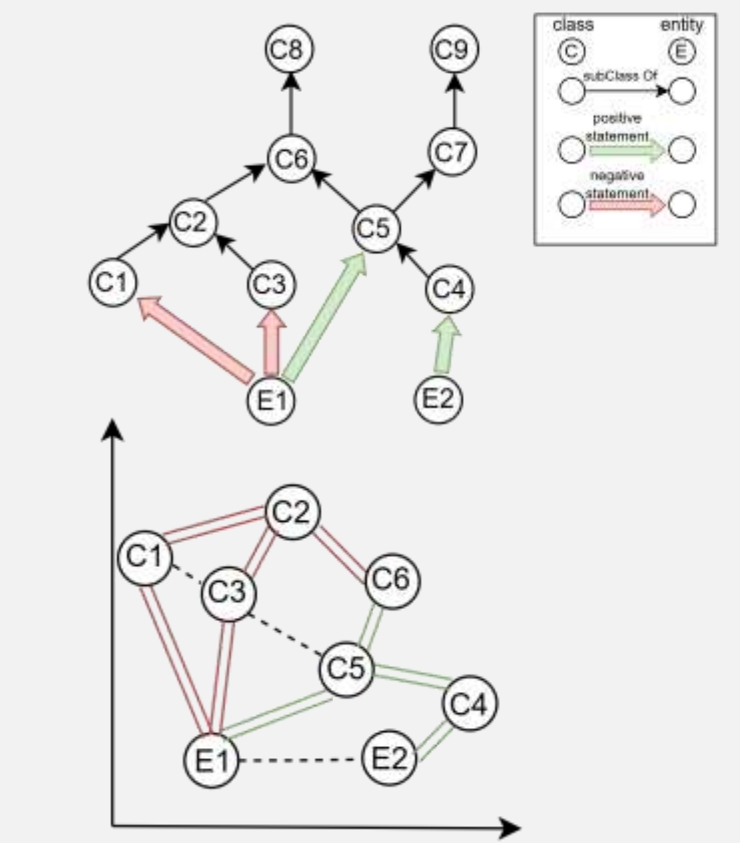
RQ 2 Does explicitly modeling contradictions and exploring this information in KG entity representations improve ML performance with contradictory facts?

RQ 3 Can external sources of knowledge (e.g. LLMs, KGs, ontologies) be explored to detect implicit contradictions?

Overview



modeling opposite relations



- Separate negative and positive relations
- Approximate negative neighbors of the same node and relation type

Preliminary study

Focus on a simpler case:

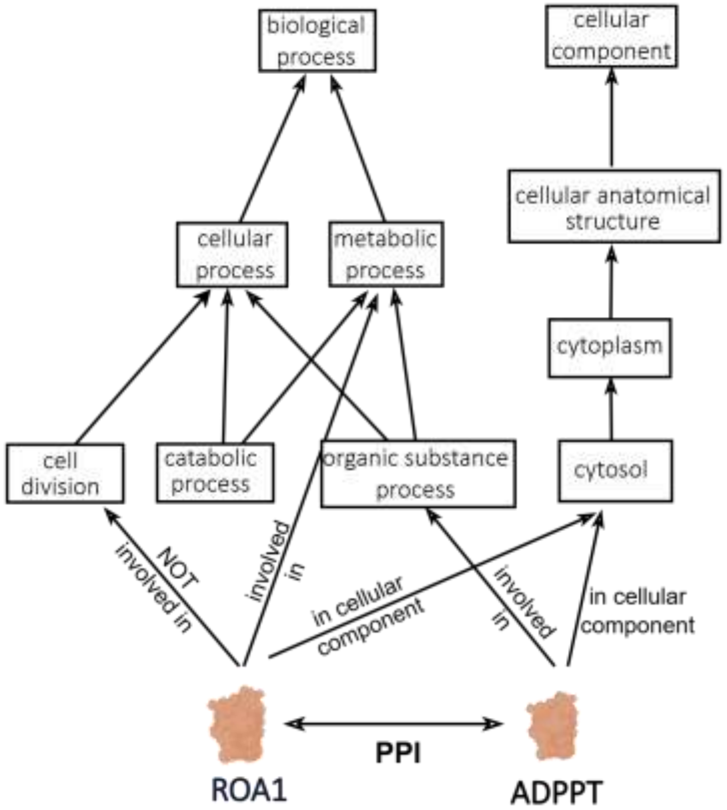
- **logical contradictions generated by negation** and on a **scientific use-case**.
- Establishing KGRL contradiction modeling baselines first.

scientific use-case:

A relevant biomedical task is the prediction of interactions between proteins (**PPIs**):

- Interactions can be modeled as a **PPI Network** (**proteins** = **nodes** and **PPIs** = **edges**);
- PPI prediction is seen as a Link Prediction/Node Pair Classification problem.

Proteins' functional aspects can help predict PPIs and they are commonly represented using biomedical ontologies (e.g. Gene Ontology).



"PPI KG" from String Protein-Protein Interaction (PPI) data and Gene Ontology

KG	# entities	# positive statements	# negative statements	"PPI" relations	Logical Contradictions*
PPI KG	12,882	74,902	69,211	215,278	8,400

Table 1. Description of the KG for the biomedical use-case for a PPI prediction task.

* detected with HermiT reasoner

Preliminary study – KGRL for PPI prediction task

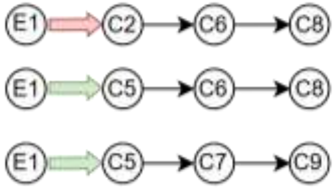
KG Representation Learning

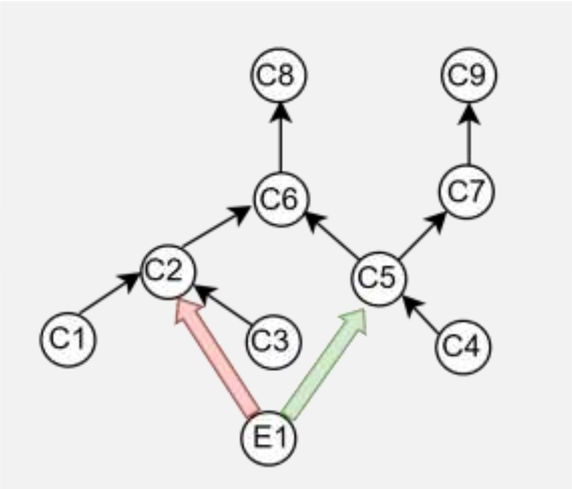
KG Embedding

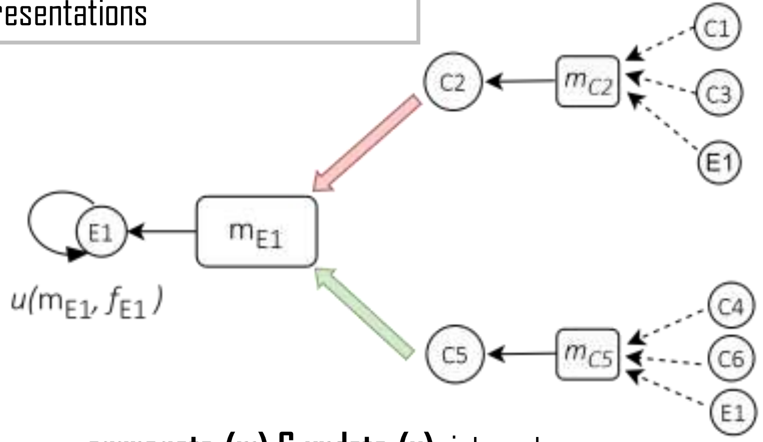
- fixed representations
- equate **proximity** with **semantic similarity**

Graph Neural –based (e.g. GNNs)

- task-specific models
- equate **proximity** with **structural similarity**
- leverage domain-specific attributes for richer representations







aggregate (m) & update (u): integrate neighbor messages and update node representation

Preliminary study – KGRL for PPI prediction task

Common KGRL embeds entities based on proximity (**proximity = similarity** assumption)

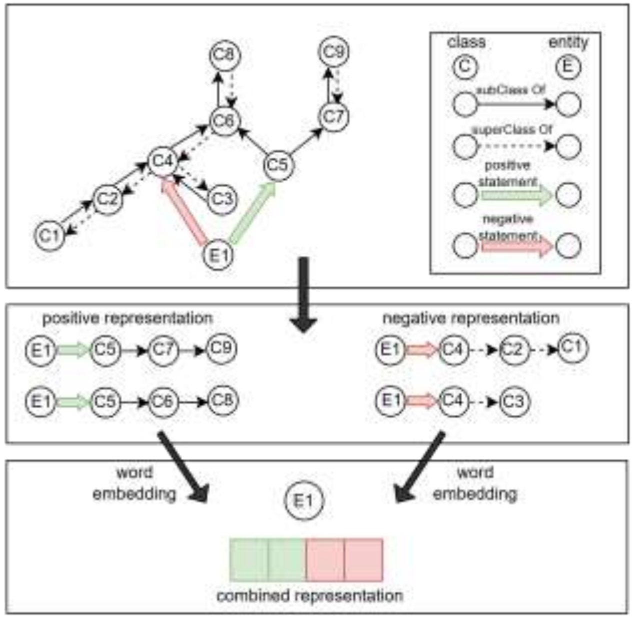
fails to model **opposite-valued relations** differently

Recent developments:

" TrueWalks (...) walk-generation method able to (...) differentiate between positive and negative statements but also take into account the semantic implications of negation"

Sousa et al. (2023)

class inheritance

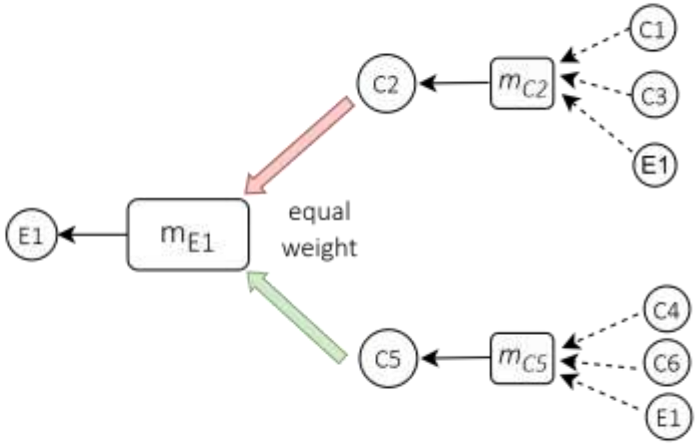


- KG Embedding –based
- Ignores modeling concerns between negative relations

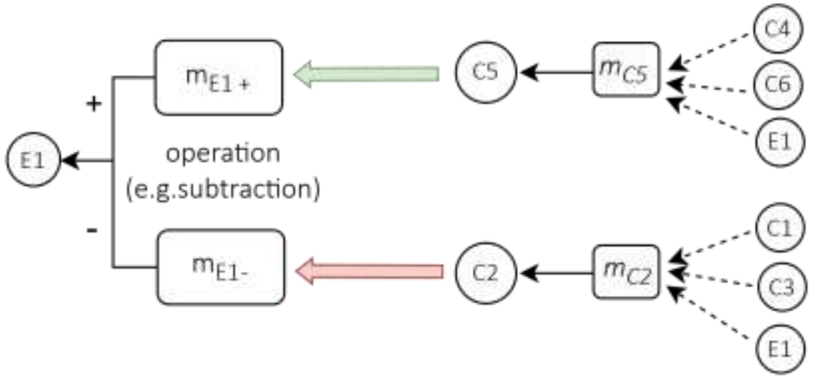
How to tackle the modeling challenge on a graph neural network?

Preliminary study – KGRL Related Work

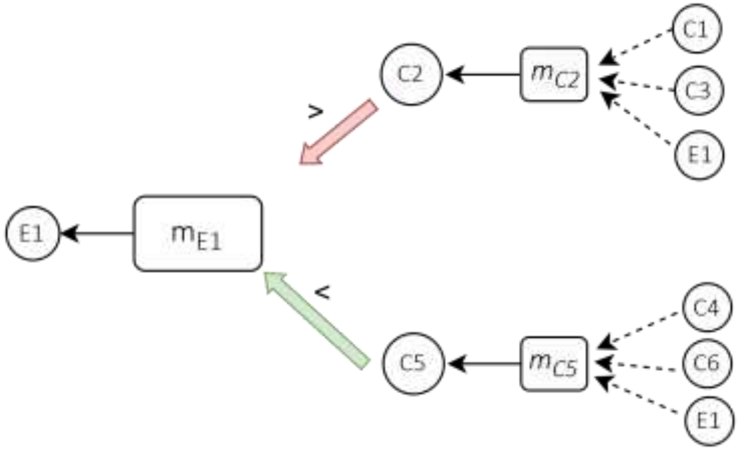
Efforts in graph neural network –based modeling of opposite relations so far:



Heterogeneous
- model graphs with multiple node and edge types.
- type-specific transformations/attentions to learn distinct semantics for each relation.
- **no explicit modeling of opposition, just different relations**.



Signed
- model graphs **with binary signals** – positive and negative edges.
- split message-passing between positive and negative neighbors.
- no modeling of different semantic relations besides positive – negative.



Heterophilic
- model graphs where connected nodes are often dissimilar.
- combined aggregation with mix and/or filter strategies over different levels of propagation
- **reduce/ignore the contribution of dissimilar neighbors during message-passing**

Preliminary study – Baseline experiments

- **PPI KG with negation-based explicit contradictions** detected by logical inference (HermiT)
- **Two KG variants**: Presence/Absence of contradictions
- **Three types of GNNs** for studying different levels of heterogeneity in a KG:
  - Homogeneous – GNNs
  - Binary Signals (**+** and **-** relations) – Bi-GNNs
  - Heterogeneous – H-GNNs
- 5-fold cross-validation of a **binary classification** of potential pairs of proteins (PPIs).

KG	Data description
Full	All statements
w/o LC	Logical Contradictions removed

Table 2. Description of Knowledge Graph variants of KG PPI.

Preliminary study – Baseline Results

GNN	KG	Accuracy	F-measure	Precision	Recall	Roc-Auc
GCN	Full	**0.6502**	**0.6071**	0.9440	**0.3193**	**0.9023**
	w/o LC	0.6484	0.6041	**0.9444**	0.3153	0.8998
BiGCN	Full	0.6107	0.5451	**0.9574**	0.2316	**0.9002**
	w/o LC	**0.6275**	**0.5719**	0.9546	**0.2675**	0.8976
HGCN	Full	0.6656	0.6294	0.9402	0.3536	**0.8916**
	w/o LC	**0.6808**	**0.6501**	**0.9441**	**0.3844**	0.8888

GNN	KG	Accuracy	F-measure	Precision	Recall	Roc-Auc
GAT	Full	**0.5721**	**0.4795**	**0.9375**	**0.1546**	**0.6913**
	w/o LC	0.5690	0.4760	0.9310	0.1491	0.6778
BiGAT	Full	**0.5673**	**0.4688**	**0.9452**	**0.1438**	**0.6642**
	w/o LC	0.5498	0.4420	0.8853	0.1140	0.6255
HGAT	Full	**0.5662**	**0.4702**	0.9386	0.1414	**0.7030**
	w/o LC	0.5661	**0.4702**	**0.9391**	**0.1415**	0.6985

Tables 4 & 5. Results for experiments over PPI KG Full and w/o LC with GCN and GAT variants considering varying levels of heterogeneity. Best values for each model in **bold.** Best values overall underlined.

KG	Data description
Full	All statements
w/o LC	Logical Contradictions removed

Table 2. Description of Knowledge Graph variants of KG PPI.

Performance Comparison

- Generally improved by including contradictions

Including contradictions does not degrade model performance, instead can improve it!

Contradictions may contain valuable but under-utilized information.

Improving the modeling of contradictions might boost predictive performance.

Preliminary Study: Key Takeaways

Validated: Contradictions can be included in KGRL without negative impact to task-specific ML performance.

Explanation: Inclusion of negative information may enhance predictions despite introducing contradictions.

Open Challenges remaining:

- Detecting other types of explicit contradictions.
- Detecting implicit contradictions (requires extra semantics, not just logic).
- Explicit modeling of the semantic value of contradictions.

Next steps: Toward the Full Thesis Scope

- Expansion of study to other types of contradictions, explicit and implicit, with existence proven in general domain (use-cases with Wikidata and ConceptNet).
- Development of approaches that combine symbolic and sub-symbolic systems for complex contradiction detection.
- Development of modeling mechanisms that integrate opposite relations differently during GNNs' message-passing.

Expected Contributions & Future Remarks

- Formal classification of contradictions in KGs and definition of implicit contradictions.
- First effort in modeling contradictions into entity representations.
- Study of how different contradictions can be captured by KGRL algorithms to improve the robustness, reliability, transparency and predictive performance of ML over KGs.
- Systematic evaluation of the approaches over several benchmark datasets and on different domain tasks.

- Potential to improve the expressiveness of KG representations and their applications — thereby contributing to the future themes of **polyvocal AI** and the **mitigation of misconceptions**, both of which depend on the richness and context that KGs can provide.

Acknowledgements & Team

LISBON SEMANTIC DATA lab

Supervisor: Catia Pesquita

Contacts:

lbalbi@lasige.di.fc.ul.pt

https://github.com/lbalbi

This work was supported by FCT through the fellowship 2024.01208.BD, and the LASIGE Research Unit, ref. UID/00408/2025.

It was also partially supported by the KATY project which has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101017453.

LASIGE

fct Fundação para a Ciência e a Tecnologia

Ciências ULisboa

KATCHY Clinical Knowledge for Humanity