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

LASIGE

U LISBOA

Ciências ULisboa

Laura Balbi

Supervised by Cátia Pesquita

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**

Real-world knowledge inherently includes ambiguity and contradictions can **reflect multiple views** valid in their own context/under a given assumption

e.g. Wikidata [1] entry for Jerusalem:

< Jerusalem, wiki: capitalOf, **S. Palestine** >
— disputed by USA & Israel

< Jerusalem, wiki:capitalOf, **Israel** >
— disputed by United Natio

[1]

Contradictions in Knowledge Graphs

KGs with explicit contradictions mirror real-world

that should be modeled and exploited for ML!

But this would mean tackling…

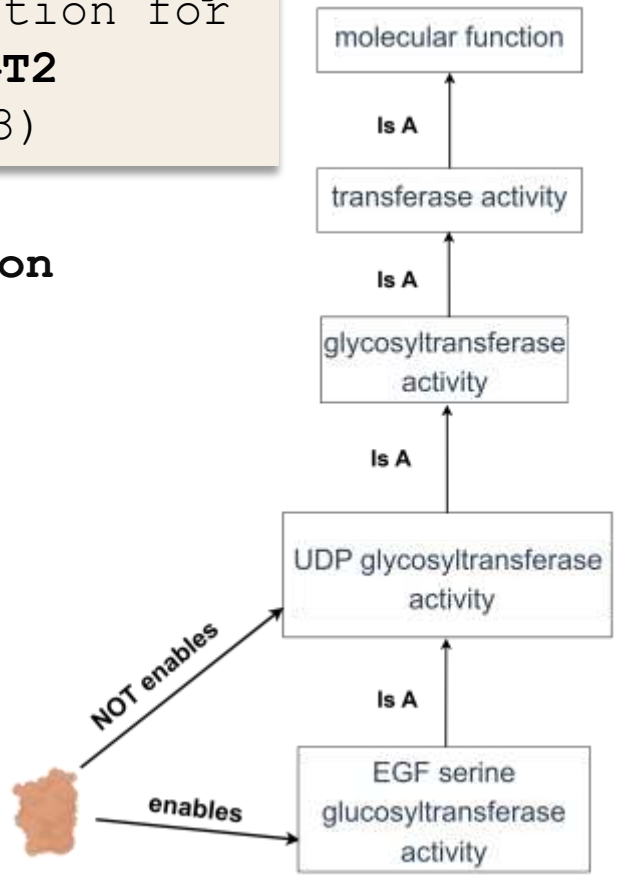
Common assumptions

- **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 n

Common solution

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

General domain – statements from diverse and conflicting viewpoints

Jerusalem, wiki:CapitalOf, **S. Palestine** – "disputed by USA & Israel"

Jerusalem, wiki: CapitalOf, Israel – "disputed by UN"

Implicit contradiction

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

Possible solution

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

Hypothesis

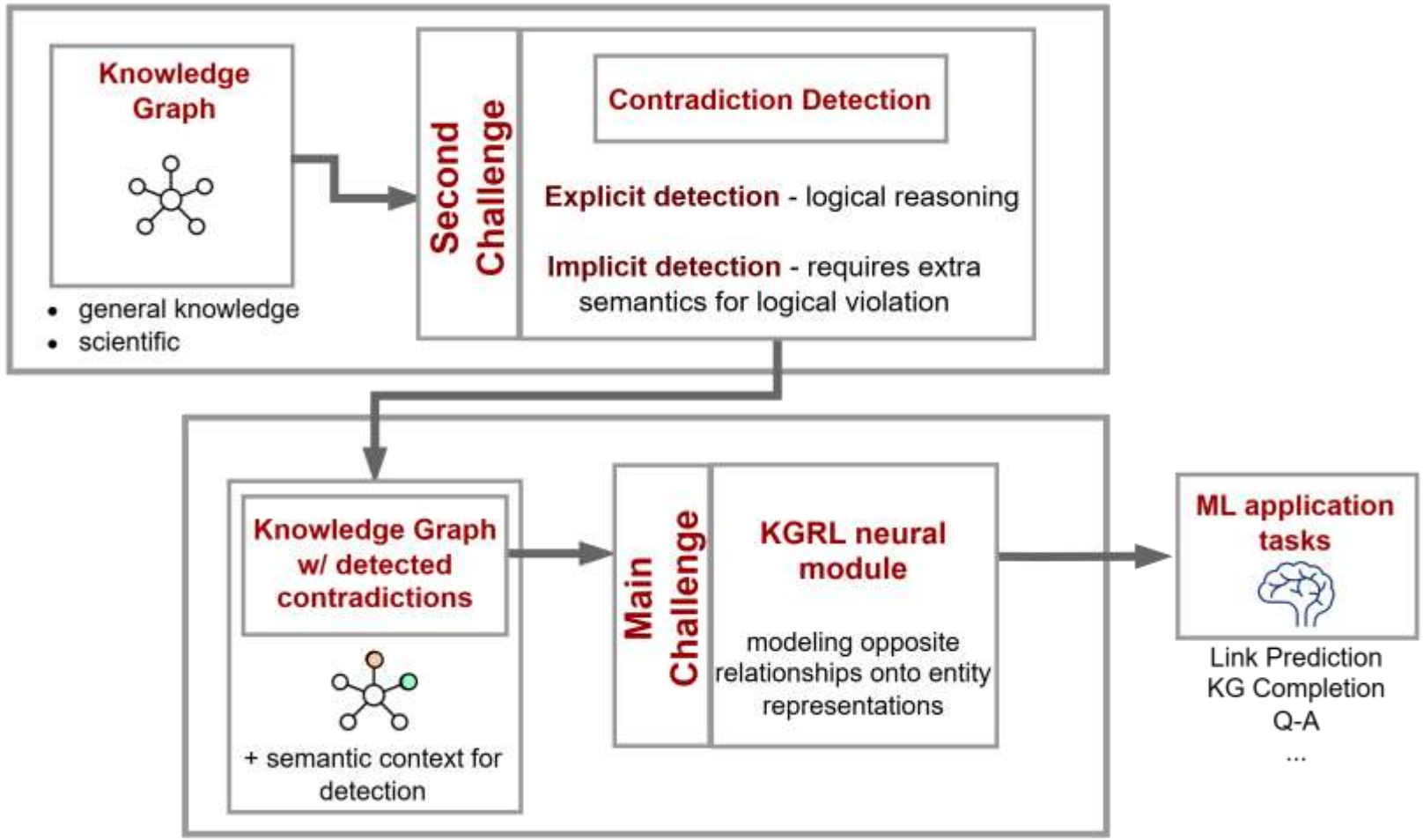
Recap

- 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.

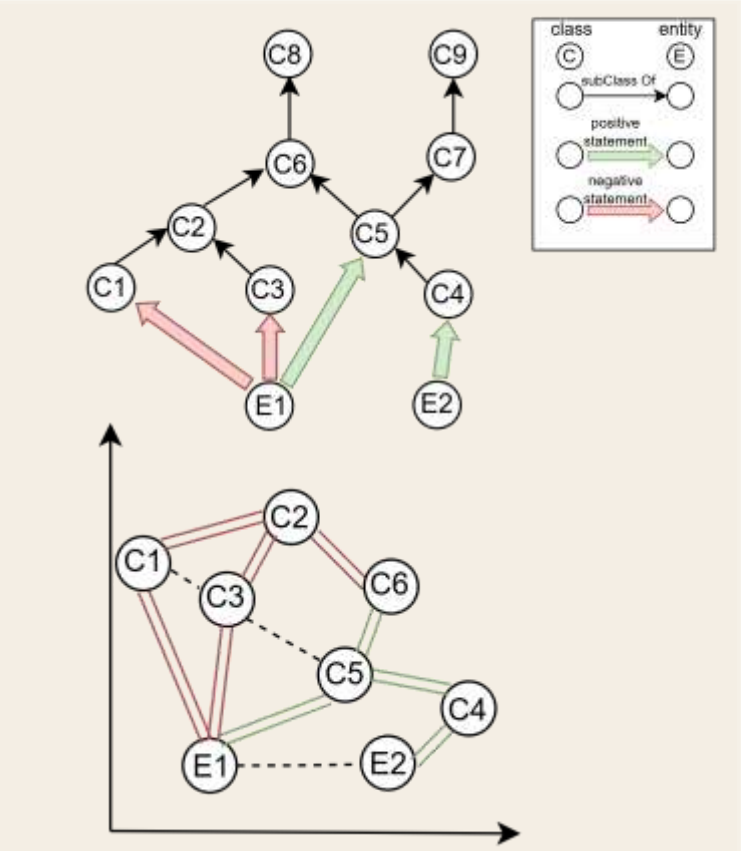
Hypothesis

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.

Methodology



modeling opposite relations



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

Acknowledgements & Team

LISBON SEMANTIC DATA lab

Supervisor: Catia Pesquita

Contacts:

lbalbi@lasige.di.fc.ul.pt

https://github.com/lbalbi

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

It is 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

KATY Clinical Knowledge for Humanity