Neural-symbolic Knowledge Representation and Reasoning

Jiaoyan Chen
Lecturer in Department of Computer Science
University of Manchester

ESSAI 2024 Athens

This course is

- introductory
- aimed at general computer scientist

- taught by
  - Uli Sattler - days 1-2
  - **Jiaoyan Chen - days 3-5**

(Hiking in Egina, Greece, 11/2023)

- explores combination/integration/collaboration of
  - Symbolic &
  - **Neural**
    - approaches to knowledge representation, reasoning, ML, …

Overview of this course

Day	Topic	Concepts	Technologies
1	Knowledge Graphs	parsing/serialisation, queries, schemas, validation & reasoning	RDF(S), SPARQL, SHACL,
2	Ontologies	Facts & background knowledge, entailments, reasoning & materialisation	OWL, OWL API, Owlready, Protégé
3	Knowledge Graph Embeddings	Classis Es, variants, inductive inference, literal-aware Es, incremental Es, application	TransE, TransH, TransR, GCN, R-GCN, OntoZSL, RMPI
4	Ontology Embeddings	Geometric embeddings, literal-aware OEs, faithfulness, evaluation & applications	ELEm, Box^2EL, OWL2Vec*, LogMap-ML, ZSL, mOWL
5	Language Models & KR, Discussion & Outlook	LM for KR, ontology & KG for LLM	BERTMap, BERTSubs, DeepOnto, ICON, BLINKOut, GraphRAG

Day 4 Ontology Embeddings

MANCHESTER 1824

What is ontology?

Knowledge representation of a domain (e.g., concepts/classes, instances/entities, properties, and logical relationships)

$$\mathcal{T} = \{\text{Father} \sqsubseteq \text{Parent} \sqcap \text{Male},\ \text{Mother} \sqsubseteq \text{Parent} \sqcap \text{Female},$$
$$\text{Child} \sqsubseteq \exists \text{hasParent}.\text{Father},\ \text{Child} \sqsubseteq \exists \text{hasParent}.\text{Mother},$$
$$\text{hasParent} \sqsubseteq \text{relatedTo}\}$$
$$\mathcal{A} = \{\text{Father}(\text{Alex}), \text{Child}(\text{Bob}), \text{hasParent}(\text{Bob}, \text{Alex})\}$$

A toy ontology on a family

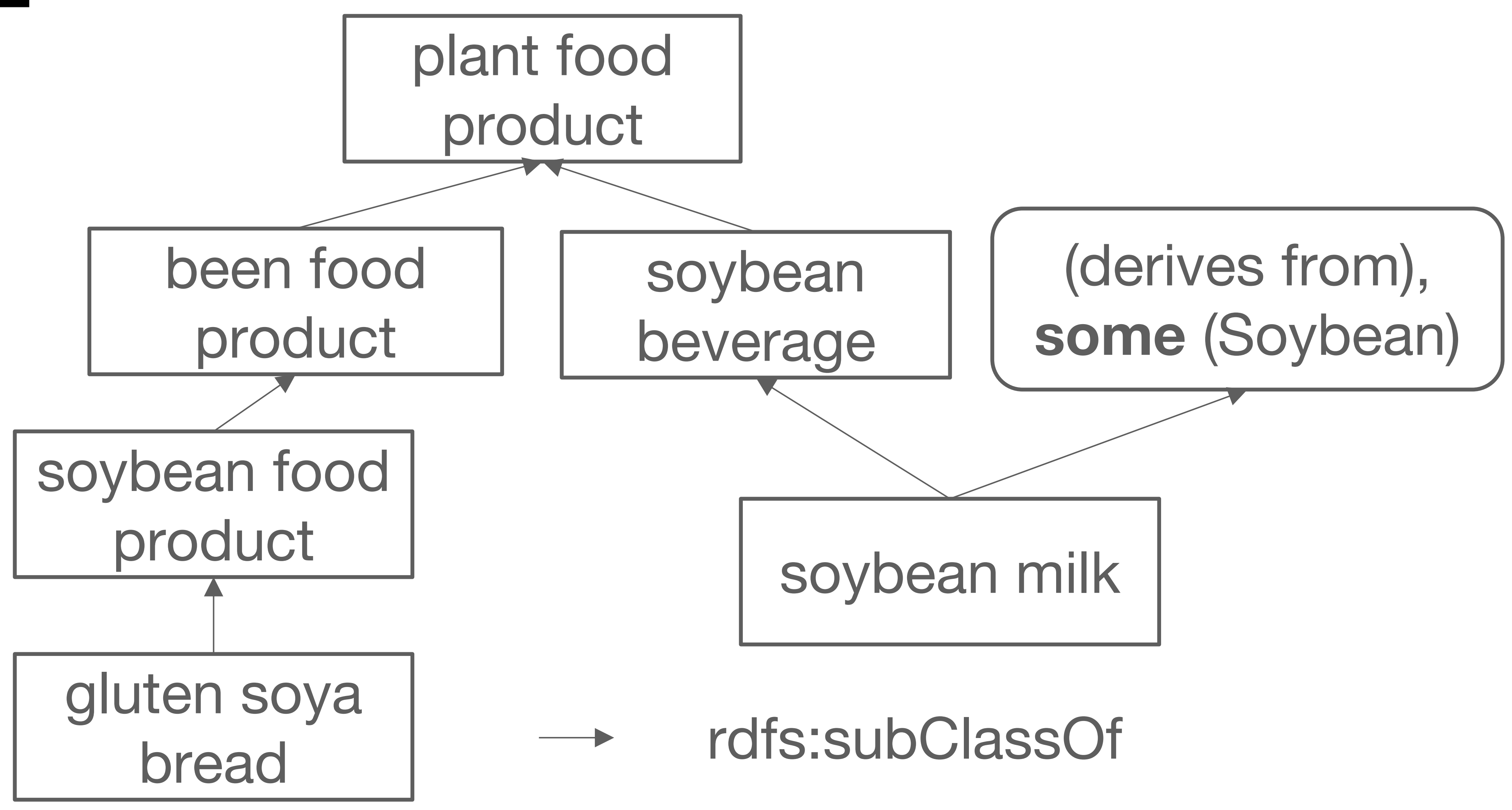
- Formal
- Explicit
- Shared

Ontology Languages

- RDF, RDFS
- **Web Ontology Language** (OWL)
  - Schema and logical relationships (domain knowledge)
  - Taxonomies and vocabularies

$$
\begin{aligned}
\mathcal{T} = \{&\mathsf{Father} \sqsubseteq \mathsf{Parent} \sqcap \mathsf{Male},\ \mathsf{Mother} \sqsubseteq \mathsf{Parent} \sqcap \mathsf{Female},\\
&\mathsf{Child} \sqsubseteq \exists \mathsf{hasParent}.\mathsf{Father},\ \mathsf{Child} \sqsubseteq \exists \mathsf{hasParent}.\mathsf{Mother},\\
&\mathsf{hasParent} \sqsubseteq \mathsf{relatedTo}\}\\
\mathcal{A} = \{&\mathsf{Father}(\mathsf{Alex}), \mathsf{Child}(\mathsf{Bob}), \mathsf{hasParent}(\mathsf{Bob}, \mathsf{Alex})\}
\end{aligned}
$$

The toy ontology on a family



An example from the food ontology FoodOn

Target Ontologies to Embed

- Simple ontologies with e.g., taxonomies
- Ontologies in RDFS or OWL (with Description Logic)
- Ontologies with literals
- Ontologies with large-scale KGs

Are KG embeddings applicable?

- Yes, because ontologies can be transformed into an RDF graph
  - E.g., W3C OWL to RDF Graph mapping
  - E.g., projection rules

- However, they are NOT for ontology embeddings as the semantics of (OWL) ontologies is much more complex
  - They introduce intermediate blank nodes or lose much semantics
  - How to distinguish the semantics of concepts and instances? How to separate the TAox and ABox?
  - How to model the logical relationships between concepts?

Description Logic $\mathcal{EL}^{++}$

Complex concepts are recursively defined as:

$$\top \mid \bot \mid A \mid C \sqcap D \mid \exists r.C \mid \{a\}$$

With role composition and inclusion:

$$r_1 \circ \cdots \circ r_k \sqsubseteq r$$

A widely used segment of Description Logic due to its good balance between expressivity and reasoning complexity (polynomial)

Corresponding to **OWL 2 EL** profile

Description Logic $\mathcal{EL}^{++}$

$$
\begin{aligned}
\mathcal{T} = \{&\text{Father} \sqsubseteq \text{Parent} \sqcap \text{Male},\ \text{Mother} \sqsubseteq \text{Parent} \sqcap \text{Female},\\
&\text{Child} \sqsubseteq \exists \text{hasParent}.\text{Father},\ \text{Child} \sqsubseteq \exists \text{hasParent}.\text{Mother},\\
&\text{hasParent} \sqsubseteq \text{relatedTo}\}\\
\mathcal{A} = \{&\text{Father}(\text{Alex}), \text{Child}(\text{Bob}), \text{hasParent}(\text{Bob}, \text{Alex})\}
\end{aligned}
$$

Revisit the toy family ontology
which is of $\mathcal{EL}^{++}$

EL Embedding with Concept as Ball

- Geometric modeling:
  - Each concept by a high dimension ball (a center and a radius)
  - Each instance by a point
  - Each binary relation (role) by a translation vector

- An embedding η is composed of two mapping functions (f_η, r_η)
  - f_η: $\boldsymbol{C} \cup \boldsymbol{R} \rightarrow \mathbb{R}^n$, r_η : $\boldsymbol{C} \rightarrow \mathbb{R}$ ($\boldsymbol{C}$ denotes the concept set, $\boldsymbol{R}$ denotes the relation set)

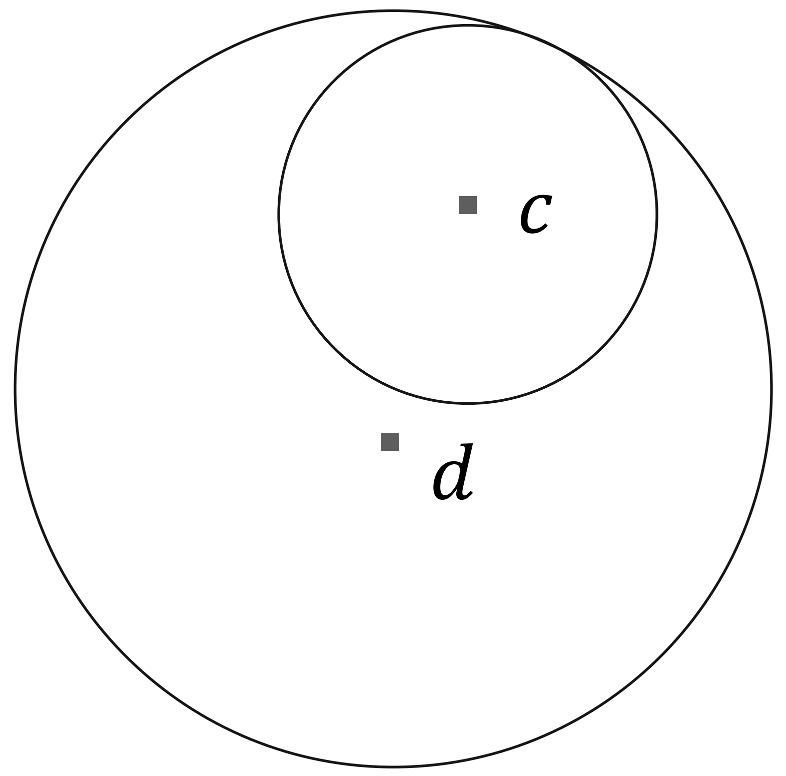
Kulmanov, Maxat, et al. "El embeddings: Geometric construction of models for the description logic el++." IJCAI 2019.

EL Embedding with Concept as Ball

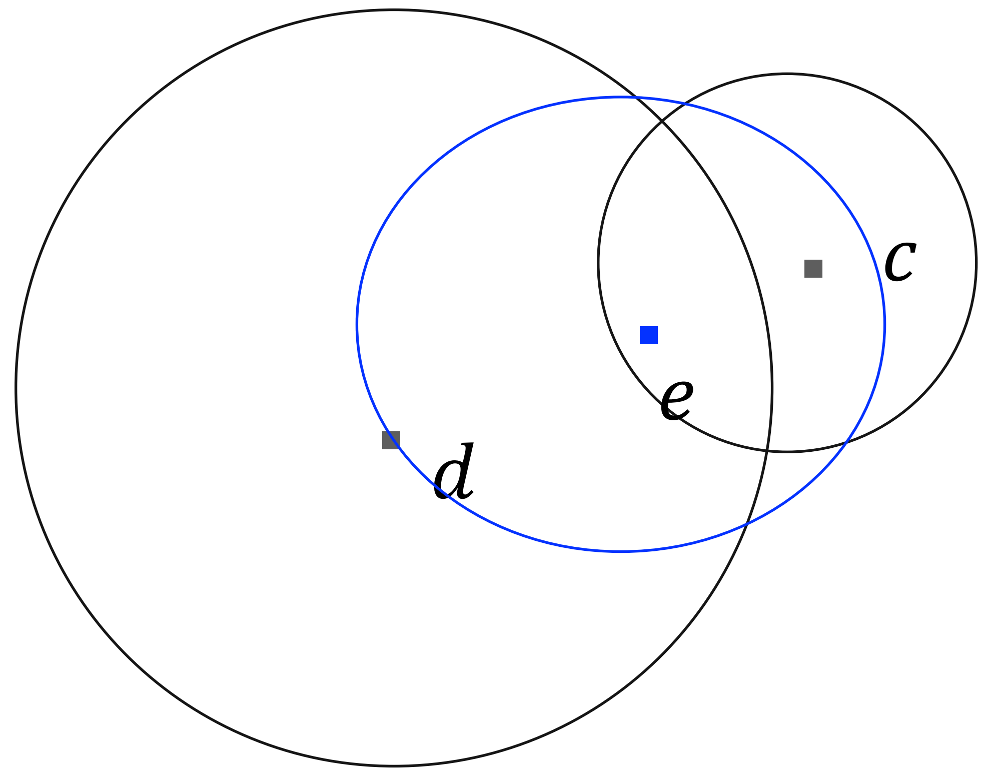
- Training
  - EL ontology is normalized into axioms of basic forms
    - $C \sqsubseteq D$, $C \sqcap D \sqsubseteq E$, $C \sqsubseteq \exists R.D$, $\exists R.C \sqsubseteq D$, $C \sqsubseteq \bot$, $C \sqcap D \sqsubseteq \bot$, $\exists R.C \sqsubseteq \bot$
  - Define score function for each form
  - Define loss with positive and negative axioms
  - Learning by stochastic gradient descent.

EL Embedding with Concept as Ball

$$loss_{C \sqsubseteq D}(c, d) = \max(0, \|f_\eta(c) - f_\eta(d)\| + r_\eta(c) - r_\eta(d) - \gamma) + |\,\|f_\eta(c)\| - 1| + |\,\|f_\eta(d)\| - 1| \quad (1)$$



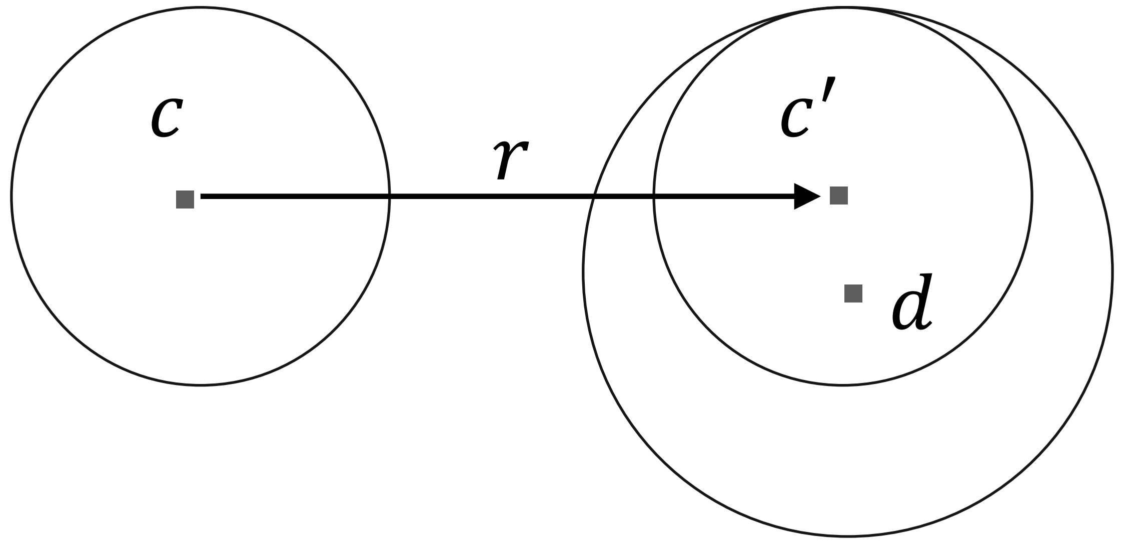
$$loss_{C \sqcap D \sqsubseteq E}(c, d, e) = \max(0, \|f_\eta(c) - f_\eta(d)\| - r_\eta(c) - r_\eta(d) - \gamma) + \max(0, \|f_\eta(c) - f_\eta(e)\| - r_\eta(e) - \gamma) + \max(0, \|f_\eta(d) - f_\eta(e)\| - r_\eta(e) - \gamma) + \max(0, \min(r_\eta(c), r_\eta(d)) - r_\eta(e) - \gamma) + |\,\|f_\eta(c)\| - 1| + |\,\|f_\eta(d)\| - 1| + |\,\|f_\eta(e)\| - 1| \quad (2)$$



(2) is an over approximation.
The conjunction of two balls is no longer a ball

EL Embedding with Concept as Ball

$$loss_{C \sqsubseteq \exists R.D}(c,d,r) = \max(0, \|f_\eta(c) + f_\eta(r) - f_\eta(d)\| + r_\eta(c) - r_\eta(d) - \gamma) + |\,\|f_\eta(c)\| - 1| + |\,\|f_\eta(d)\| - 1| \quad (3)$$



$$loss_{\exists R.C \sqsubseteq D}(c,d,r) = \max(0, \|f_\eta(c) - f_\eta(r) - f_\eta(d)\| + r_\eta(c) - r_\eta(d) - \gamma) + |\,\|f_\eta(c)\| - 1| + |\,\|f_\eta(d)\| - 1| \quad (4)$$

$$loss_{C \sqcap D \sqsubseteq \bot}(c,d,e) = \max(0, r_\eta(c) + r_\eta(d) - \|f_\eta(c) - f_\eta(d)\| + \gamma) \quad (5)$$
$$+ |\,\|f_\eta(c)\| - 1| + |\,\|f_\eta(d)\| - 1|$$

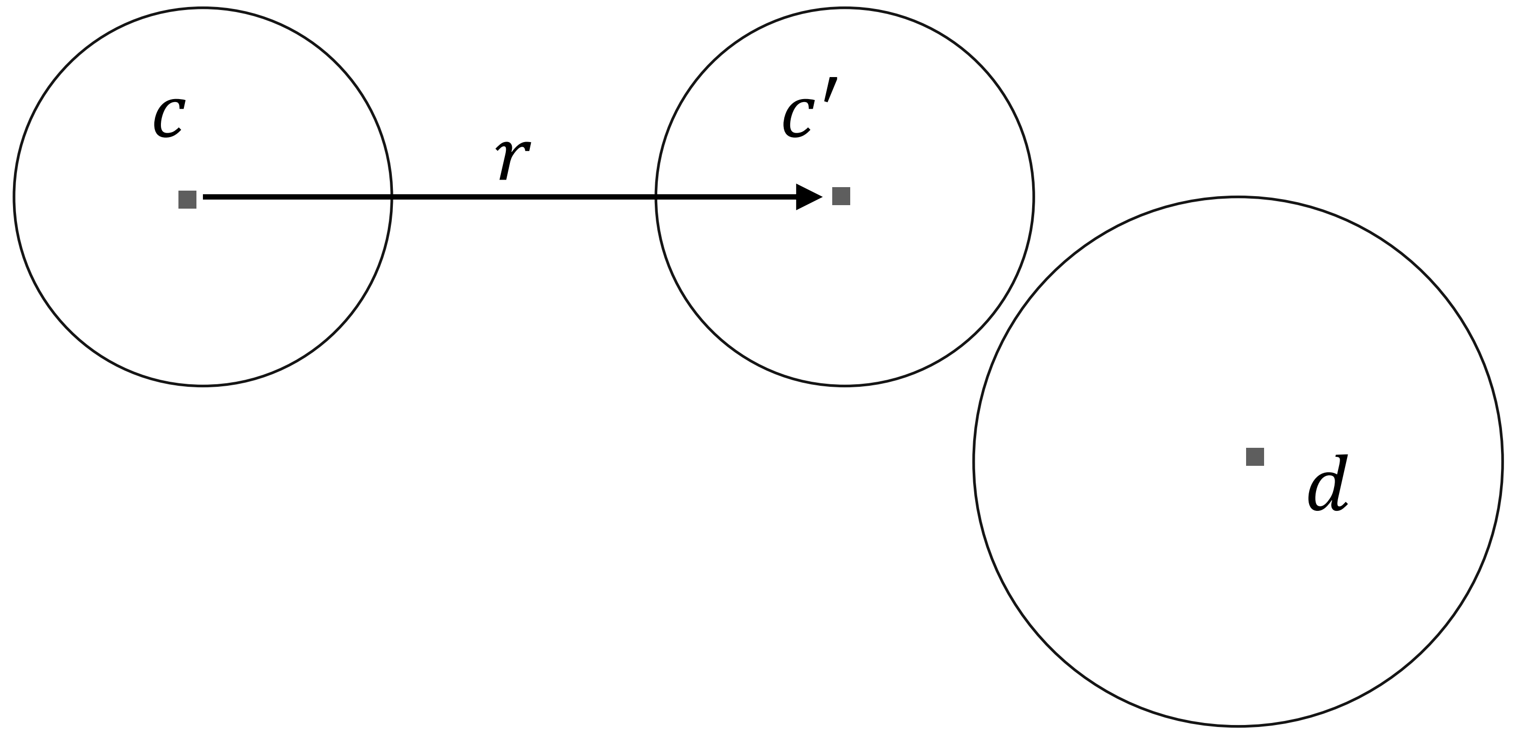
$$loss_{C \sqsubseteq \bot}(c) = r_\eta(c) \quad (6)$$

$$loss_{\exists R.C \sqsubseteq \bot}(c,r) = r_\eta(c) \quad (7)$$

EL Embedding with Concept as Ball

- Negative samples
  - Corrupt axioms of $C \sqsubseteq \exists R.D$ for negative axioms in form of $C' \not\sqsubseteq \exists R.D$ or $C \not\sqsubseteq \exists R.D'$

$$
\begin{aligned}
&loss_{C \not\sqsubseteq \exists R.D}(c, d, r) = \\
&\max(0, r_\eta(c) + r_\eta(d) - \|f_\eta(c) + f_\eta(r) - f_\eta(d)\| + \gamma) \\
&+ |\, \|f_\eta(c)\| - 1| + |\, \|f_\eta(d)\| - 1|
\end{aligned}
\tag{8}
$$



Box^{2}EL: EL Embedding with Concept as Box

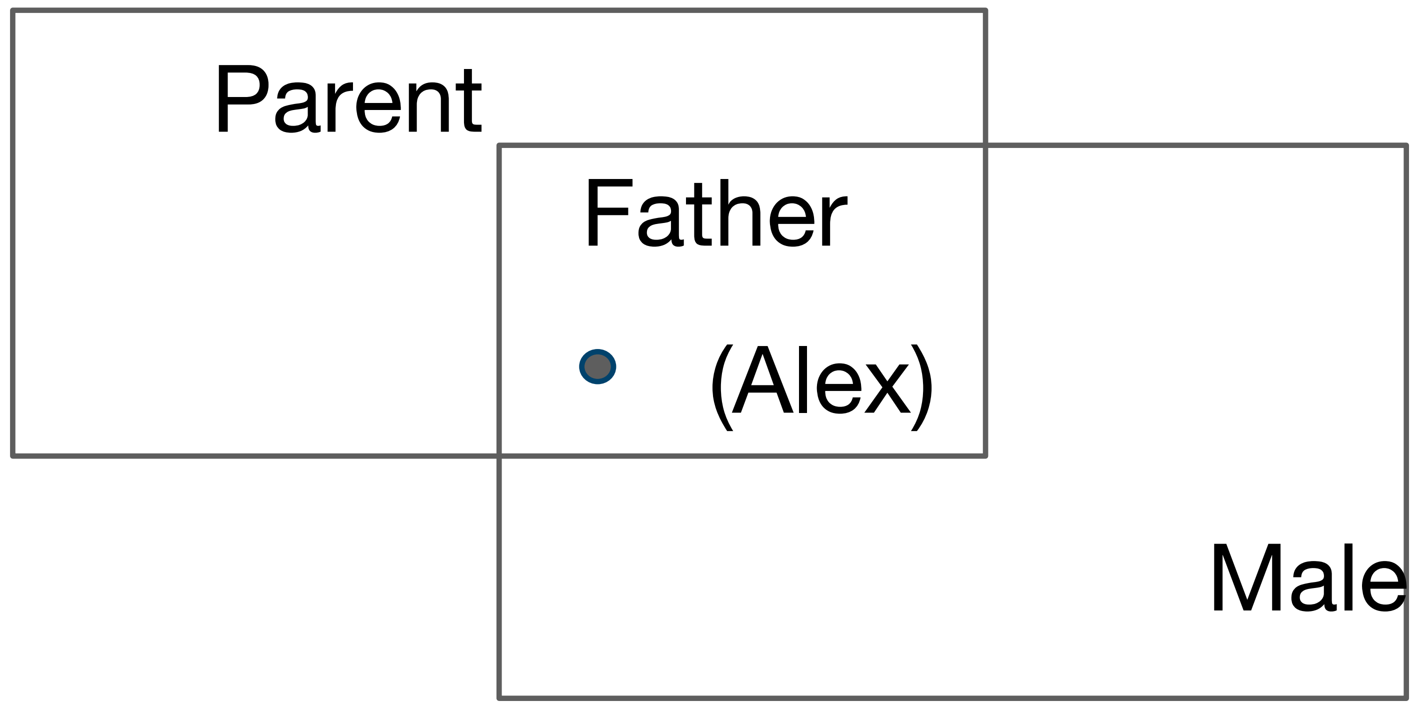
- Concept as ball: the intersection of two balls is no longer a ball
- Simple vector based translation cannot model one-to-many, many-to-one and many-to-many relations

- Box^{2}EL aims to address the two limitations
  - Concept as box
  - Bump vector for modeling concept relationship

Mathias, J., et al. "Dual box embeddings for the description logic EL++." *Proceedings of the ACM on Web Conference 2024*. 2024.

Box^{2}EL: EL Embedding with Concept as Box

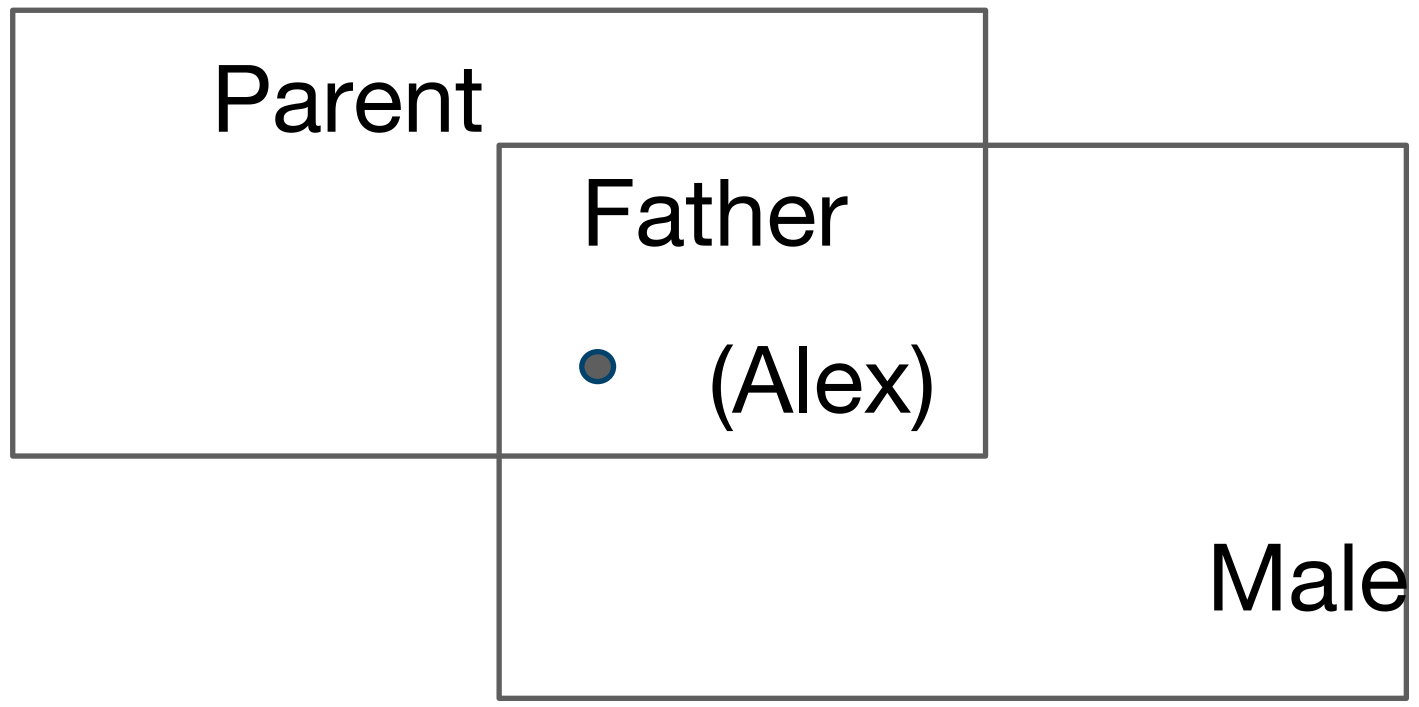
Concept: Box
Instance: Point



How about $C \sqsubseteq \exists r.D$?

Box^2EL: EL Embedding with Concept as Box

Concept: Box
Instance: Point



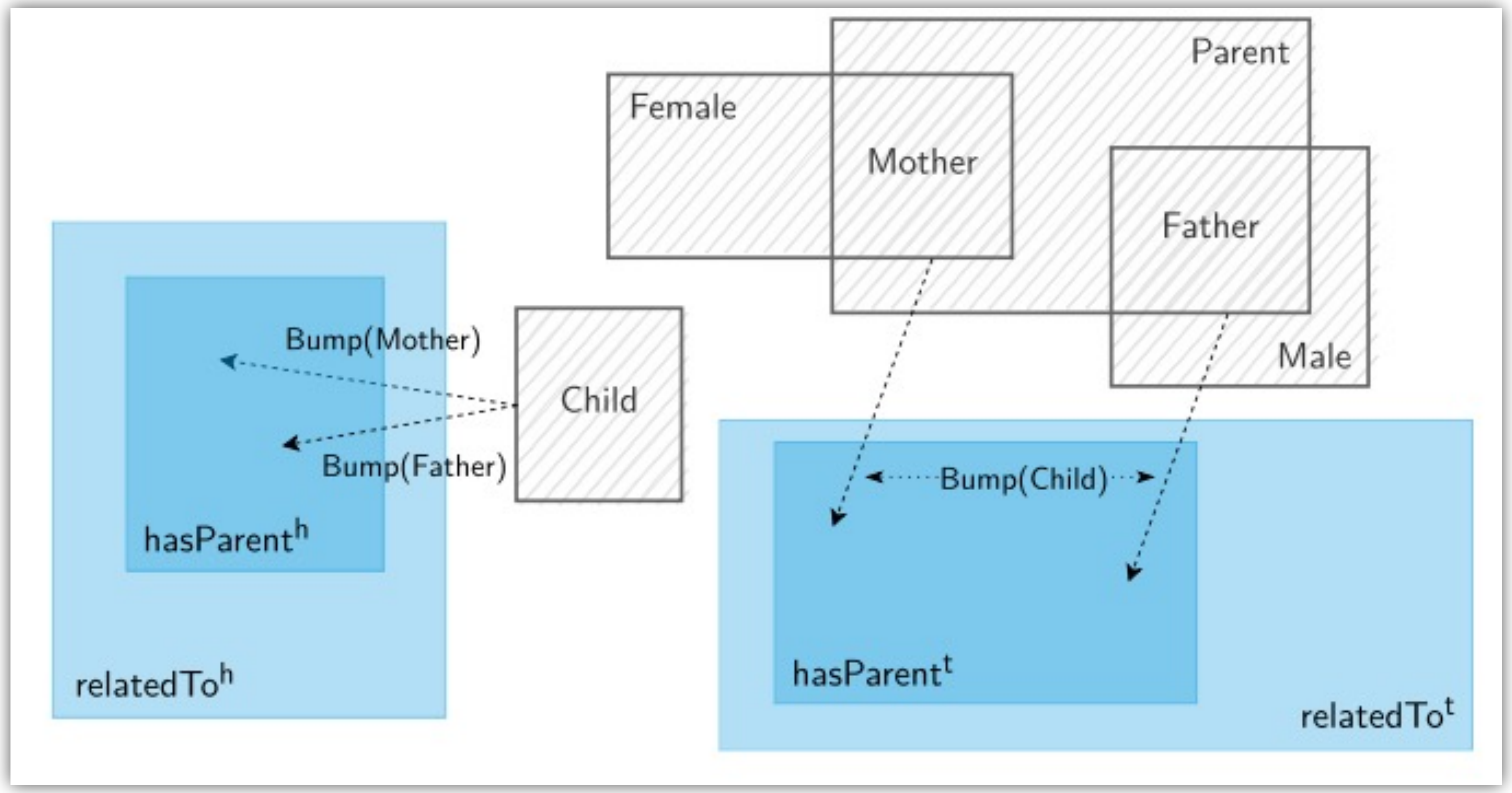
Relation: Head Box & Tail Box

$C \sqsubseteq \exists r.D$: Box(C) + Bump(D) ⊆ Head(r)
Box(D) + Bump(C) ⊆ Tail(r)

Box^{2}EL: EL Embedding with Concept as Box

$$\mathcal{T} = \{\text{Father} \sqsubseteq \text{Parent} \sqcap \text{Male},\ \text{Mother} \sqsubseteq \text{Parent} \sqcap \text{Female},$$
$$\text{Child} \sqsubseteq \exists\text{hasParent.Father},\ \text{Child} \sqsubseteq \exists\text{hasParent.Mother},$$
$$\text{hasParent} \sqsubseteq \text{relatedTo}\}$$
$$\mathcal{A} = \{\text{Father(Alex)}, \text{Child(Bob)}, \text{hasParent(Bob, Alex)}\}$$

A toy family ontology



Concept: Box
Instance: Point

Relation: Head Box & Tail Box

$C \sqsubseteq \exists r.D$: Box(C) + Bump(D) ⊆ Head(r)
Box(D) + Bump(C) ⊆ Tail(r)

Box^{2}EL: EL Embedding with Concept as Box

- ABox axioms are transformed to equivalent TBox axioms by **nominals**

$$
\begin{aligned}
C(a) &\rightsquigarrow \{a\} \sqsubseteq C \\
r(a,b) &\rightsquigarrow \{a\} \sqsubseteq \exists r.\{b\}
\end{aligned}
$$

Box^2EL: EL Embedding with Concept as Box

- The following normalized axioms are considered as positive:

NF1: $C \sqsubseteq D$
NF2: $C \sqcap D \sqsubseteq E$
NF3: $C \sqsubseteq \exists r.D$
NF4: $\exists r.C \sqsubseteq D$
NF5: $C \sqcap D \sqsubseteq \bot$
NF6: $\mathrm{r} \sqsubseteq s$
NF7: $r_1 \circ r_2 \sqsubseteq s$

- The axioms of NF3 are corrupted for negative axioms: $C' \not\sqsubseteq \exists R.D$ or $\mathrm{C} \not\sqsubseteq \exists R.D'$

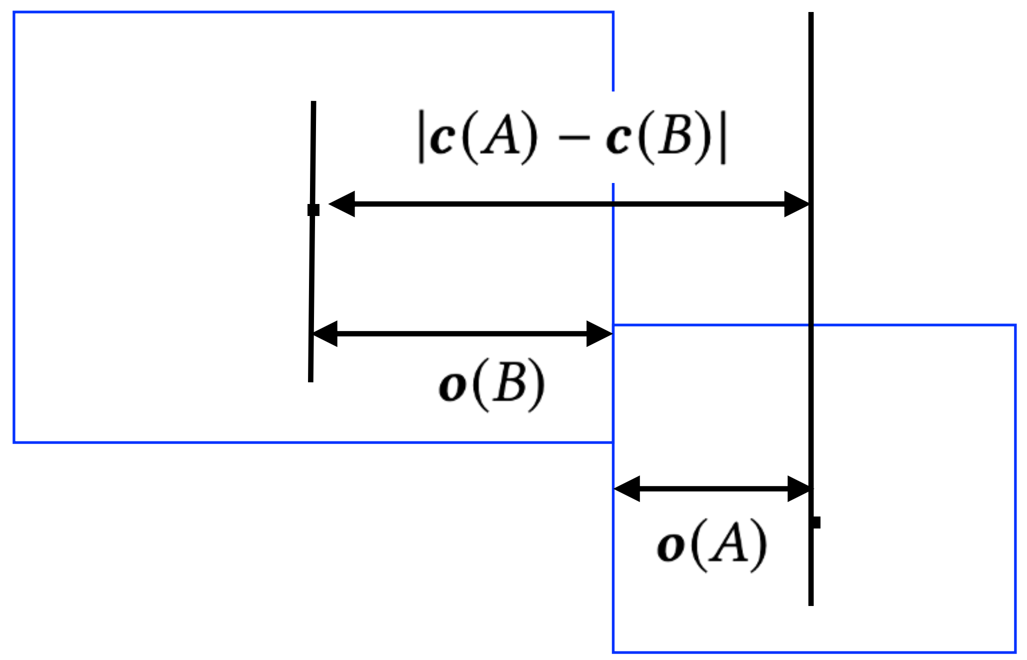
Box^{2}EL: EL Embedding with Concept as Box

- Element-wise distance of two boxes:

$$\boldsymbol{d}(A, B) = |\boldsymbol{c}(A) - \boldsymbol{c}(B)| - \boldsymbol{o}(A) - \boldsymbol{o}(B)$$

$\boldsymbol{c}(\cdot)$: center vector
$\boldsymbol{o}(\cdot)$: offset vector

($\boldsymbol{d}(A, B) \geq 0$: disjointed; otherwise, overlapped)



Demonstration on one dimension (horizontal)

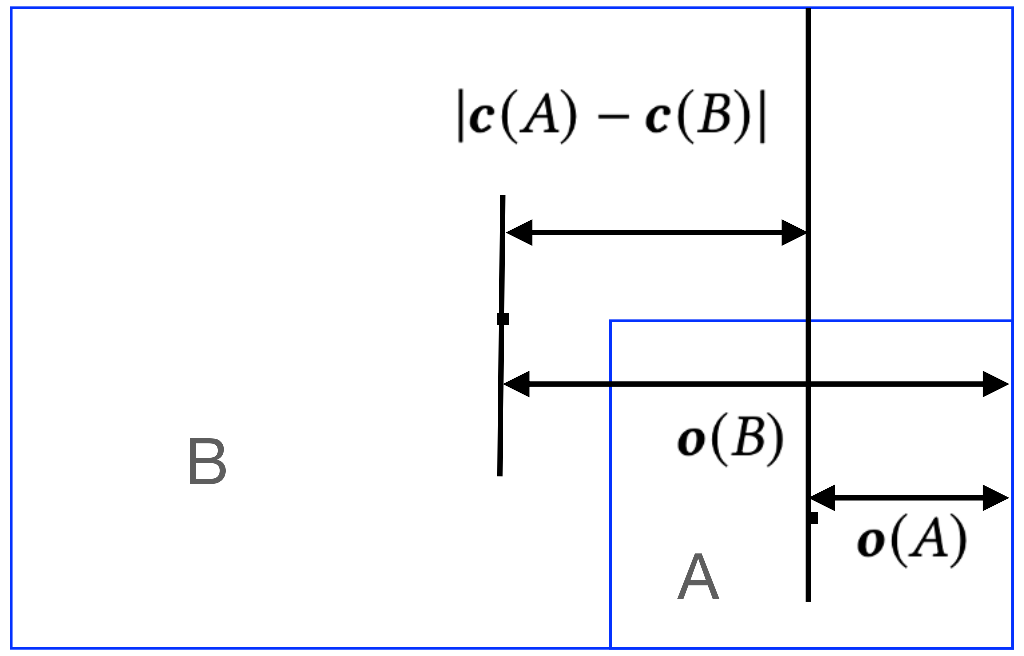
Box²EL: EL Embedding with Concept as Box

- Box containment:

$$|\boldsymbol{c}(A) - \boldsymbol{c}(B)| - \big(\boldsymbol{o}(B) - \boldsymbol{o}(A)\big) - \gamma$$

$$\mathcal{L}_{\subseteq}(A, B) = \begin{cases} \|\max\{\mathbf{0},\ \boldsymbol{d}(A, B) + 2\boldsymbol{o}(A) - \gamma\}\| & \text{if } B \neq \emptyset \\ \max\{0, \boldsymbol{o}(A)_1 + 1\} & \text{otherwise.} \end{cases}$$

($\mathcal{L}_{\subseteq}(A, B)$ <= 0: contain; otherwise, not contain)



Demonstration on one dimension (horizontal)

Box^{2}EL: EL Embedding with Concept as Box

- Losses

NF1: $C \sqsubseteq D$

NF2: $C \sqcap D \sqsubseteq E$

NF3: $C \sqsubseteq \exists r.D$

NF4: $\exists r.C \sqsubseteq D$

NF5: $C \sqcap D \sqsubseteq \perp$

NF6/7: $\mathrm{r} \sqsubseteq s$ or $r_1 \circ r_2 \sqsubseteq s$

$$\mathcal{L}_1(C, D) = \mathcal{L}_{\subseteq}(\mathrm{Box}(C), \mathrm{Box}(D))$$

$$\mathcal{L}_2(C, D, E) = \mathcal{L}_{\subseteq}\Big(\mathrm{Box}(C) \cap \mathrm{Box}(D), \mathrm{Box}(E)\Big)$$

$$\mathcal{L}_3(C, r, D) = \frac{1}{2}\Big(\mathcal{L}_{\subseteq}(\mathrm{Box}(C) + \mathrm{Bump}(D), \mathrm{Head}(r)) + \mathcal{L}_{\subseteq}(\mathrm{Box}(D) + \mathrm{Bump}(C), \mathrm{Tail}(r))\Big).$$

$$\mathcal{L}_4(r, C, D) = \mathcal{L}_{\subseteq}(\mathrm{Head}(r) - \mathrm{Bump}(C), \mathrm{Box}(D))$$

$$\mathcal{L}_5(C, D) = \|\max\{\mathbf{0}, -(\boldsymbol{d}(\mathrm{Box}(C), \mathrm{Box}(D)) + \gamma)\}\|$$

$$\mathcal{L}_6(r, s) = \frac{1}{2}\Big(\mathcal{L}_{\subseteq}(\mathrm{Head}(r), \mathrm{Head}(s)) + \mathcal{L}_{\subseteq}(\mathrm{Tail}(r), \mathrm{Tail}(s))\Big)$$

$$\mathcal{L}_7(r_1, r_2, s) = \frac{1}{2}\Big(\mathcal{L}_{\subseteq}(\mathrm{Head}(r_1), \mathrm{Head}(s)) + \mathcal{L}_{\subseteq}(\mathrm{Tail}(r_2), \mathrm{Tail}(s))\Big)$$

Faithfulness

- Briefly, an ontology embedding is *faithful* (also known as *sound*) if it preserves the structure (semantics) that it aims to preserve (i.e., all the target axioms are satisfied).
  - See the formal definition in Chen, J., et al. “Ontology Embedding: A Survey of Methods, Applications and Resources”, 2024.

THEOREM 1 (SOUNDNESS). *Let $\mathcal{O} = (\mathcal{T}, \mathcal{A})$ be an $\mathcal{EL}^{++}$ ontology. If $\gamma \leq 0$ and there exist Box^{2}EL embeddings in $\mathbb{R}^n$ such that $\mathcal{L}(\mathcal{O}) = 0$, then these embeddings are a model of $\mathcal{O}$.*

Inductive Reasoning

- General subsumption prediction for axioms of NF1 – NF4 ($C \sqsubseteq D$, $C \sqcap D \sqsubseteq E$, $C \sqsubseteq \exists r.D$, $\exists r.C \sqsubseteq D$)
  - 20% masked (10% for test, 10% for validation), 80% kept (for training)

Inductive Reasoning

- General subsumption prediction for axioms of NF1 – NF4 ($C \sqsubseteq D$, $C \sqcap D \sqsubseteq E$, $C \sqsubseteq \exists r.D$, $\exists r.C \sqsubseteq D$)
  - 20% masked (10% for test, 10% for validation), 80% kept (for training)

- Ranking-based evaluation
  - For each test axiom, generate a set of candidate predictions by replacing the atomic side of the subsumption
  - Rank all the candidate predictions by a score based on the distances of the embeddings of the concepts (and the relation)
  - Metrics: the median rank (Med), the mean reciprocal rank (MRR), the mean rank (MR), the area under the ROC curve (AUC), Hits@K

$$s(C \sqsubseteq D) = -\|\boldsymbol{c}(\mathrm{Box}(C)) - \boldsymbol{c}(\mathrm{Box}(D))\|.$$

$$s(C \sqsubseteq \exists r.D) = -\|\boldsymbol{c}(\mathrm{Box}(C)) + \mathrm{Bump}(D) - \boldsymbol{c}(\mathrm{Head}(r))\| - \|\boldsymbol{c}(\mathrm{Box}(D)) + \mathrm{Bump}(C) - \boldsymbol{c}(\mathrm{Tail}(r))\|.$$

Inductive Reasoning

	Model	H@1	H@10	H@100	Med	MRR	MR	AUC
GALEN	ELEm	0.01	0.12	0.29	1662	0.05	5153	0.78
	EmEL^{++}	0.01	0.11	0.24	2295	0.05	5486	0.76
	BoxEL	0.00	0.03	0.16	4785	0.01	7163	0.69
	ELBE	0.02	0.14	0.27	1865	0.06	5303	0.77
	Box^2EL	**0.05**	**0.20**	**0.35**	**669**	**0.10**	**4375**	**0.81**
GO	ELEm	0.03	**0.24**	0.43	272	0.09	6204	0.86
	EmEL^{++}	0.03	0.23	0.38	597	0.09	6710	0.85
	BoxEL	0.01	0.06	0.08	8443	0.03	14905	0.68
	ELBE	0.01	0.10	0.22	1838	0.04	6986	0.85
	Box^2EL	**0.04**	0.23	**0.59**	**48**	**0.10**	**3248**	**0.93**
Anatomy	ELEm	0.10	0.40	0.64	22	0.19	6464	0.94
	EmEL^{++}	0.11	0.36	0.57	36	0.19	8472	0.92
	BoxEL	0.03	0.12	0.28	1151	0.06	10916	0.90
	ELBE	0.04	0.36	0.63	29	0.15	5400	0.95
	Box^2EL	**0.16**	**0.47**	**0.70**	**13**	**0.26**	**2675**	**0.97**

← Overall results on all normal forms of testing subsumptions on three ontologies

Result of axioms of each form can be found in the paper ↓

Mathias, J., et al. "Dual box embeddings for the description logic EL++." *Proceedings of the ACM on Web Conference 2024*. 2024.

(Approximate) Deductive Reasoning

- Assess faithfulness/soundness
- Experiment setting
  - Train with all the asserted axioms of an ontology
  - Valid and test (10% & 90%) with the NF1 axioms ($C \sqsubseteq D$) that can be logically inferred by a symbolic reasoner
  - Ranking-based evaluation

(Approximate) Deductive Reasoning

	Model	H@1	H@10	H@100	Med	MRR	MR	AUC
GALEN	ELEm	0.00	0.04	0.20	1807	0.01	4405	0.81
	EmEL^{++}	0.00	0.04	0.18	2049	0.01	4634	0.81
	BoxEL	0.00	0.00	0.01	6906	0.00	7925	0.67
	ELBE	0.00	0.06	0.16	1785	0.02	3974	0.84
	Box^2EL	**0.01**	**0.09**	**0.24**	**1003**	**0.03**	**2833**	**0.88**
GO	ELEm	0.00	0.04	0.22	1629	0.02	7377	0.84
	EmEL^{++}	0.00	0.04	0.19	1346	0.01	6557	0.86
	BoxEL	0.00	0.00	0.13	1085	0.00	5359	0.88
	ELBE	0.00	0.06	0.21	935	0.02	3846	0.92
	Box^2EL	0.00	**0.08**	**0.49**	**107**	**0.04**	**1689**	**0.96**
Anatomy	ELEm	0.00	0.07	0.28	901	0.02	7958	0.93
	EmEL^{++}	0.00	0.07	0.26	1576	0.02	10976	0.90
	BoxEL	**0.01**	**0.10**	0.24	838	**0.04**	9156	0.92
	ELBE	0.00	0.08	0.32	336	0.03	2312	0.98
	Box^2EL	**0.01**	0.09	**0.44**	**152**	**0.04**	**1599**	**0.99**

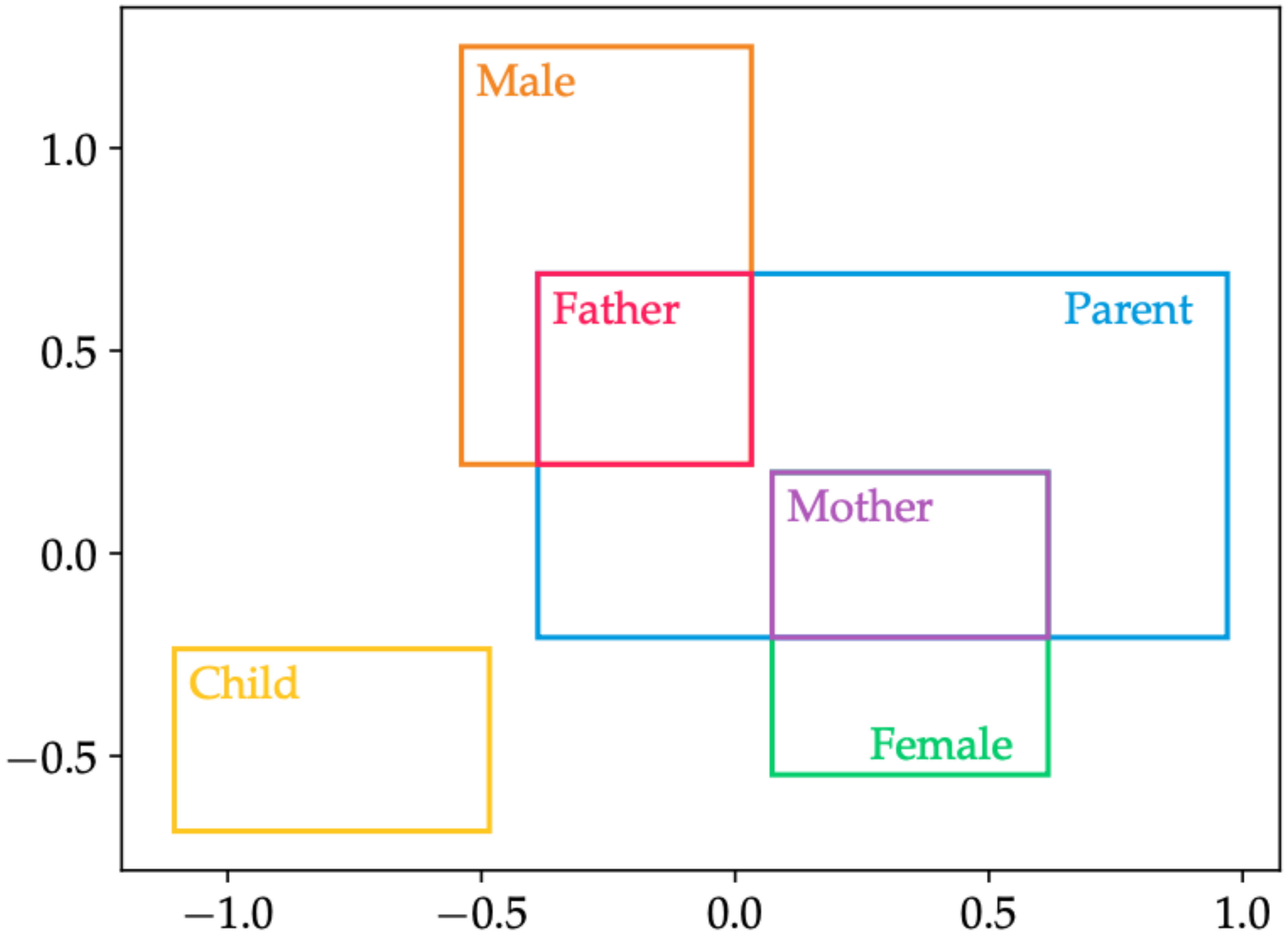
Discussion:

1. Not always 100%. Why? Faithfulness?

2. Deductive reasoning vs inductive reasoning

Visualization (Proof of Concept on the Family Ontology)

- Dimension $n = 2$
- Margin $\gamma = 0$, no negative samples, regularization length $\lambda = 1$
- An additional visualization loss to ensure the concept box volume is large enough for plotting



Link (Protein-Protein Interaction) Prediction

- OWL Ontology
  - ABox: protein-protein interactions (PPIs) from the STRING database, e.g., (P_1, interacts, P_2)
  - TBox: The Gene Ontology (GO)
- Task
  - Predict missing subsumptions in form of $\{P_1\} \sqsubseteq \exists r.\{P_2\}$
- Evaluation
  - Ranking-based metrics
  - 80%/10%/10% for train, validation and test sets

Link (Protein-Protein Interaction) Prediction

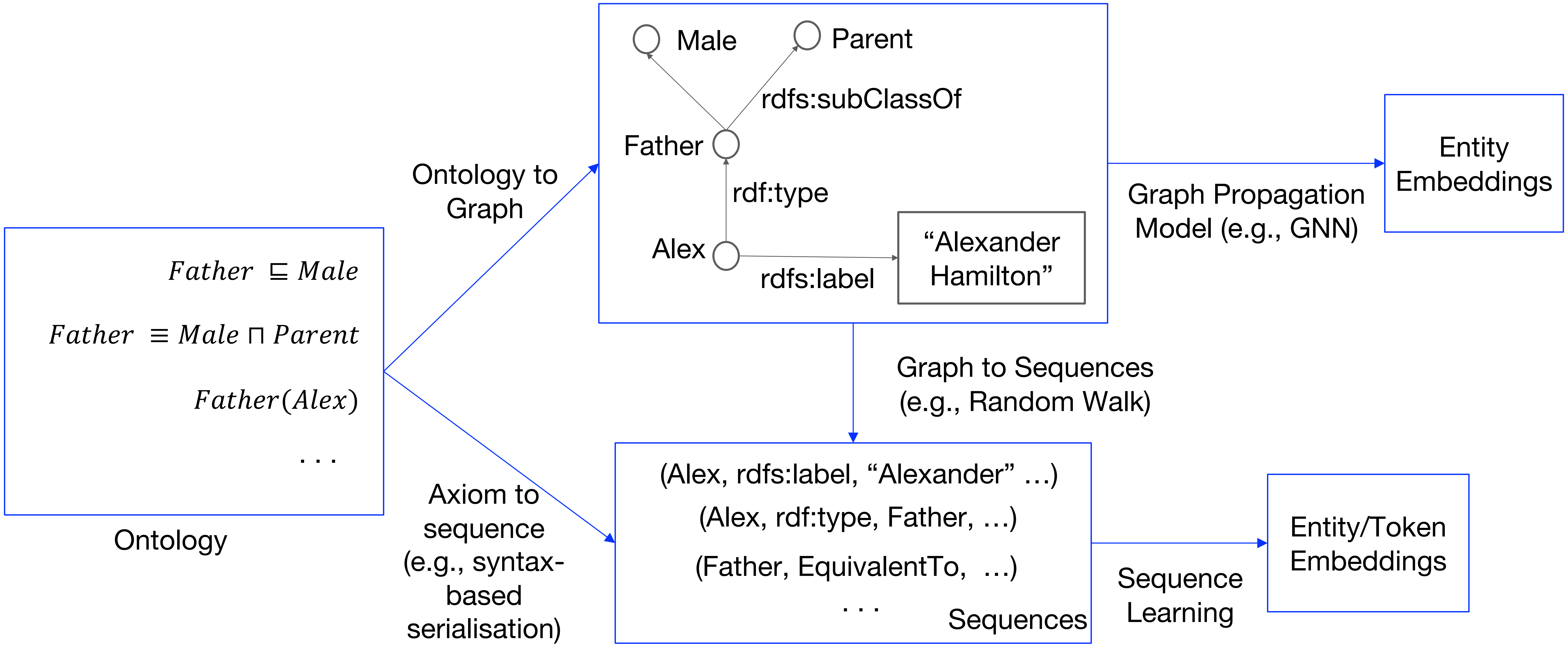
	Model	H@10	H@10 (F)	H@100	H@100 (F)	MR	MR (F)	AUC	AUC (F)
Yeast	ELEm	0.10	0.23	0.50	0.75	247	187	0.96	0.97
	EmEL^{++}	0.08	0.17	0.48	0.65	336	291	0.94	0.95
	BoxEL	0.09	0.20	0.52	0.73	423	379	0.93	0.94
	ELBE	**0.11**	0.26	0.57	0.77	201	154	0.96	0.97
	Box^{2}EL	**0.11**	**0.33**	**0.64**	**0.87**	**168**	**118**	**0.97**	**0.98**
Human	ELEm	**0.09**	0.22	0.43	0.70	658	572	0.96	0.96
	EmEL^{++}	0.04	0.13	0.38	0.56	772	700	0.95	0.95
	BoxEL	0.07	0.10	0.42	0.63	1574	1530	0.93	0.93
	ELBE	**0.09**	0.22	0.49	0.72	434	362	0.97	**0.98**
	Box^{2}EL	**0.09**	**0.28**	**0.55**	**0.83**	**343**	**269**	**0.98**	**0.98**

Paradigms for Ontology Embedding

- Geometric modeling (like Box^{2}EL)
  - Pros: interpretable; sound representation of formal semantics
  - Cons: hard to incorporate informal semantics like **textual literals**; **hard to deal with all the features of OWL**
- Sequence modeling
  - Transform axioms and literals into sentences;
  - Train word embedding (sequence learning) models
- Graph propagation
  - Transform axioms into a graph

Chen, J., et al.,"Ontology Embedding: A Survey of Methods, Applications and Resources." https://arxiv.org/abs/2406.10964.

Paradigms for Ontology Embedding



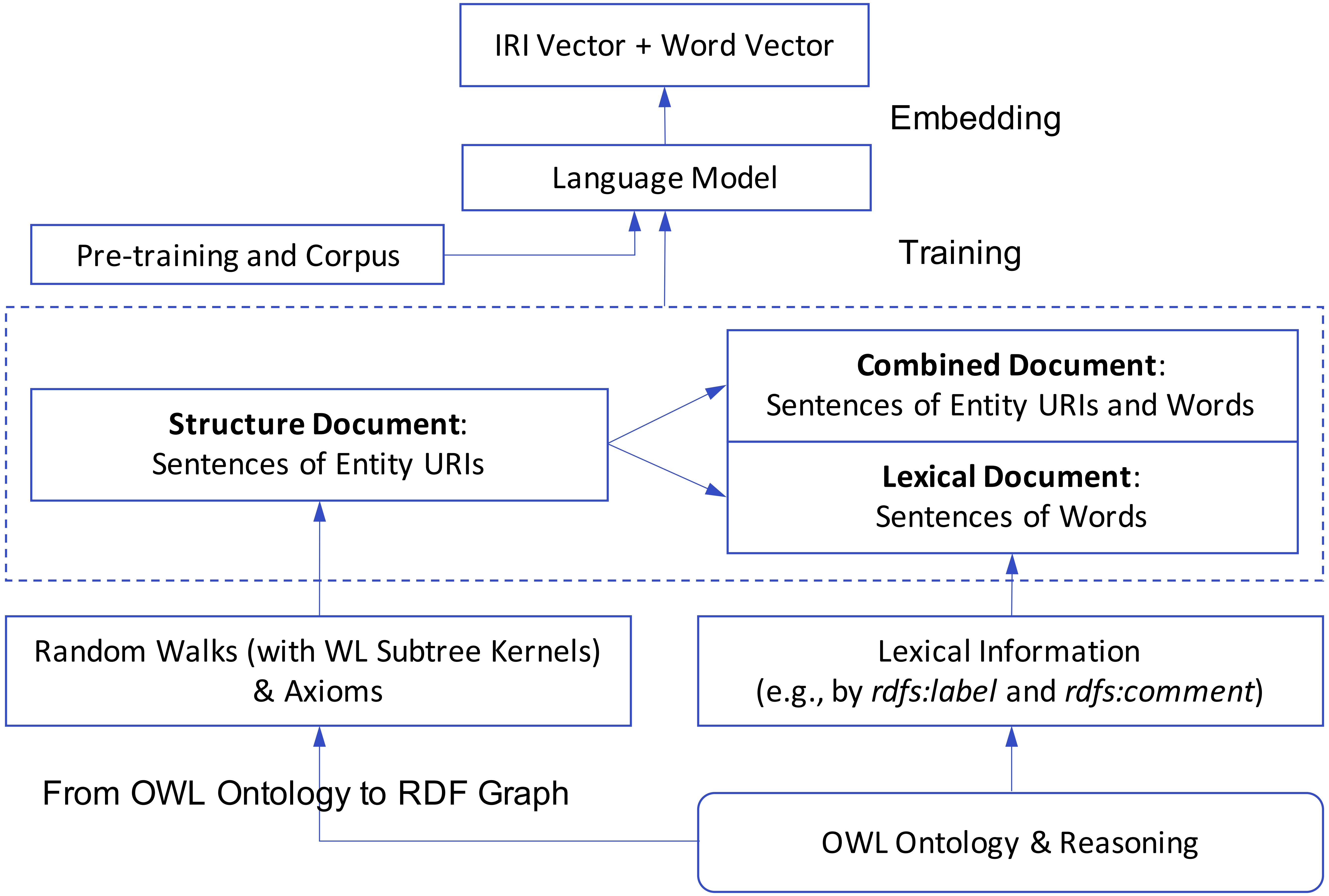
Paradigms of Sequence Learning & Graph Propagation

OWL2Vec*: OWL Ontology Embedding

- Belongs to the paradigm of sequence learning
  - Extract sequences from the ontology
  - Learn a word embedding model from the sequences
- Consider the semantics of
  - Axioms
  - Literals (text)
- Output embeddings of
  - Entities (concepts, relations and instances)
  - Tokens (words) of the text

Chen, J., et al., "Owl2vec*: Embedding of owl ontologies." *Machine Learning* 110.7 (2021): 1813-1845..

OWL2Vec*: OWL Ontology Embedding



← Pipeline

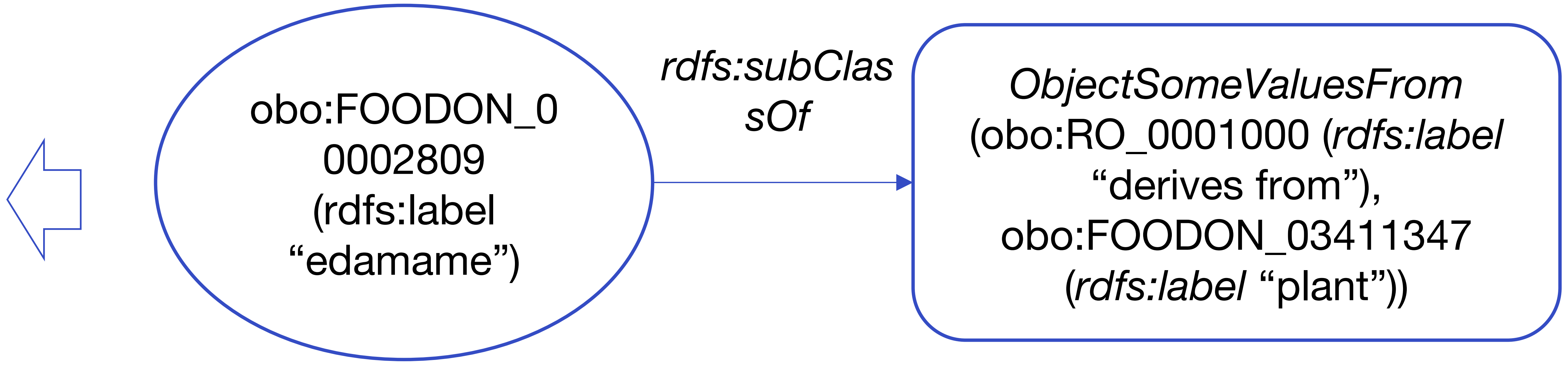
OWL2Vec*: OWL Ontology Embedding

1. From OWL Ontology to RDF Graph

- Reasoning by E.g. HermiT can be enabled

Solution #1: W3C OWL to RDF Graph Mapping

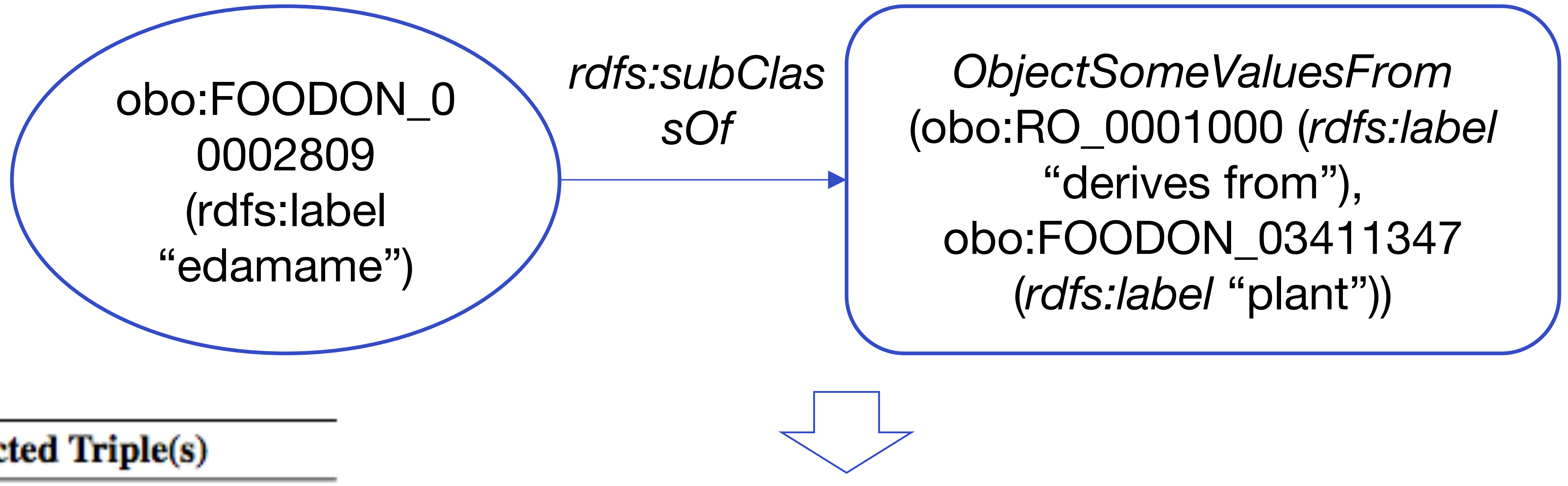
e.g.,
<obo:FOODON_00002809, rdfs:subClassOf, _:x>
<_:x, rdf:type, owl:Restriction>
<_:x, owl:OnProperty, obo:RO_0001000>
<_:x, owl:SomeValueFrom, obo:FOODON_03411347>



OWL2Vec*: OWL Ontology Embedding

1. From OWL Ontology to RDF Graph

- Reasoning by E.g. HermiT can be enabled



e.g., <obo:FOODON_00002809, rdfs:subClassOf, obo:FOODON_03411347>

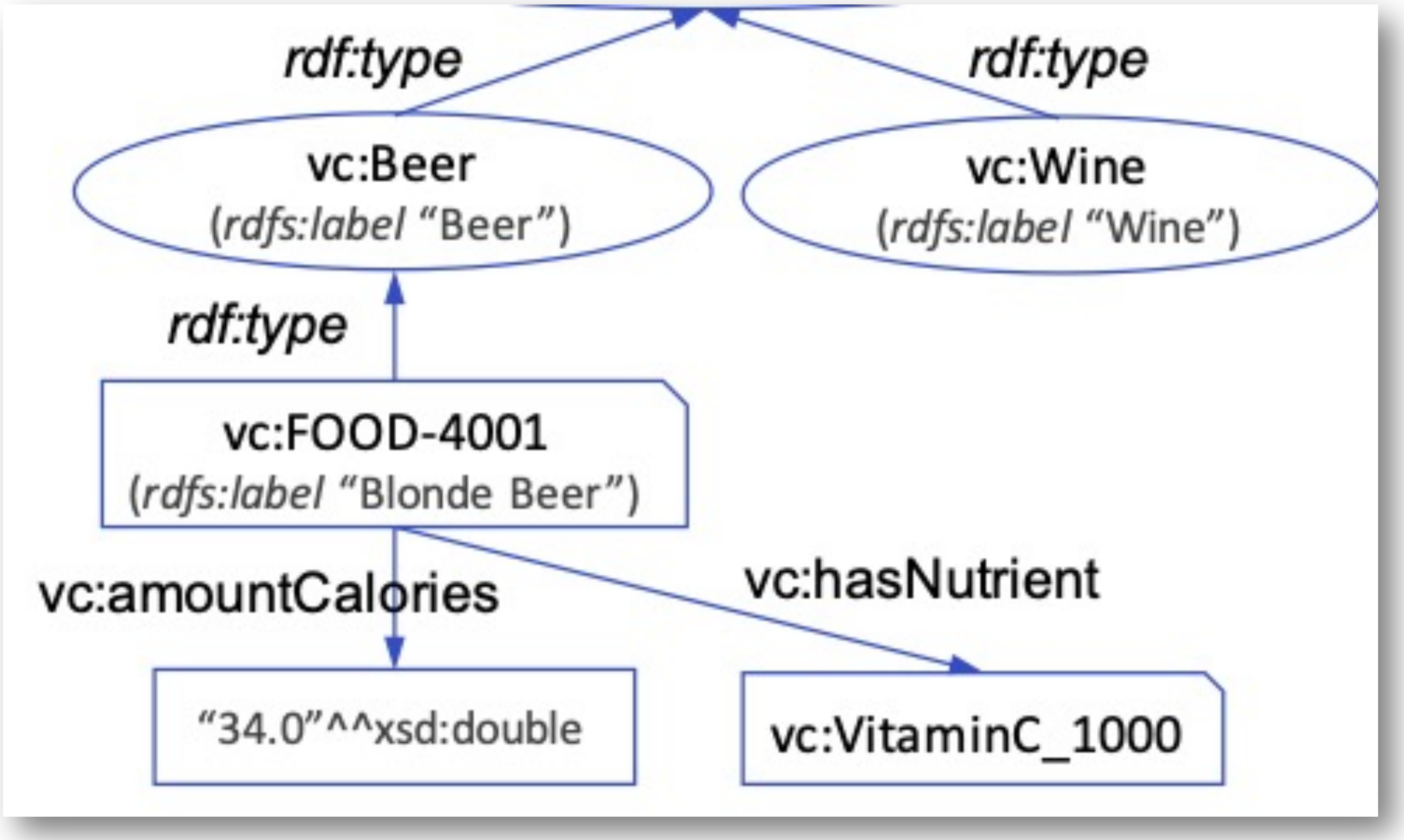
Solution #2: Projection rules

Axiom of Condition 1	Axiom or Triple(s) of Condition 2	Projected Triple(s)
$A \sqsubseteq \square r.D$ or $\square r.D \sqsubseteq A$	$D \equiv B \mid B_1 \sqcup \ldots \sqcup B_n \mid B_1 \sqcap \ldots \sqcap B_n$	$\langle A, r, B \rangle$ or $\langle A, r, B_i \rangle$ for $i \in 1, \ldots, n$
$\exists r.\top \sqsubseteq A$ (domain)	$\top \sqsubseteq \forall r.B$ (range)	
$A \sqsubseteq \exists r.\{b\}$	$B(b)$	
$r \sqsubseteq r'$	$\langle A, r', B \rangle$ has been projected	
$r' \equiv r^-$	$\langle B, r', A \rangle$ has been projected	
$s_1 \circ \ldots \circ s_n \sqsubseteq r$	$\langle A, s_1, C_1 \rangle \ldots \langle C_n, s_n, B \rangle$ have been projected	
$B \sqsubseteq A$	–	$\langle B, rdfs{:}subClassOf, A \rangle$ $\langle A, rdfs{:}subClassOf^-, B \rangle$
$A(a)$	–	$\langle a, rdf{:}type, A \rangle$ $\langle A, rdf{:}type^-, a \rangle$
$r(a, b)$	–	$\langle a, r, b \rangle$

OWL2Vec*: OWL Ontology Embedding

2. Structure document – sentences of IRIs

Solution #1: Random walk + Weisfeiler-Lehman subtree kernel



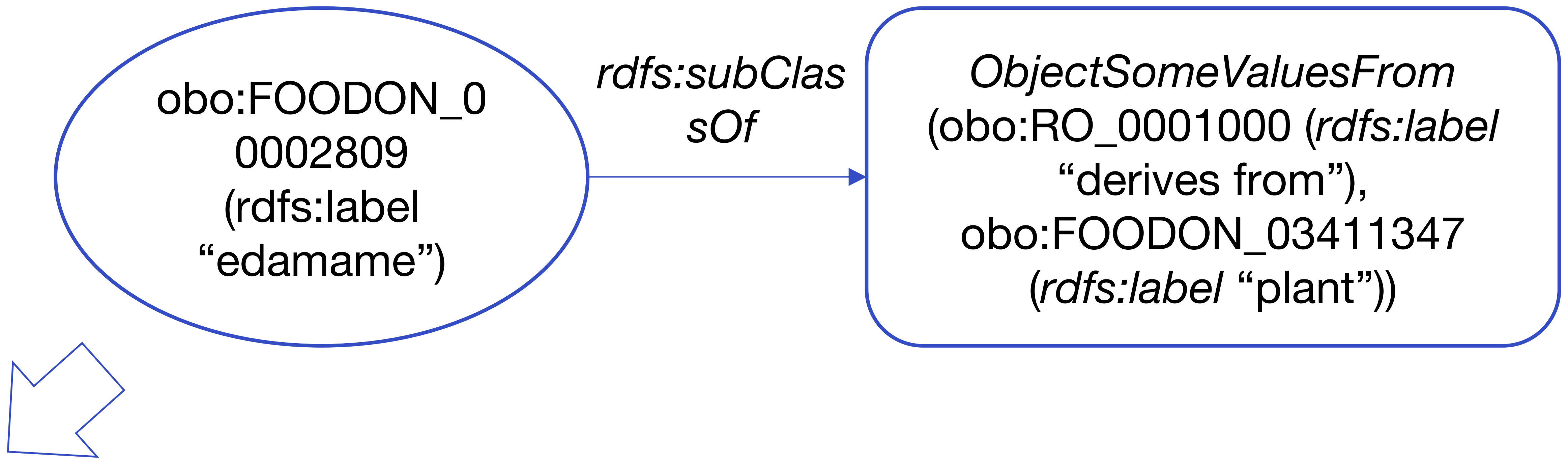
E.g.,

- (*vc:FOOD-4001*, *vc:hasNutrient*, *vc:VitaminC_100*, *vc:amountNutrient*)
- (*vc:FOOD-4001*, *rdf:type*, *kernel_id1_md5*, *rdfs:subClassOf*, *kernel_id2_md5*)

OWL2Vec*: OWL Ontology Embedding

2. Structure document – sentences of IRIs

Solution #1: Axiom serisation



E.g., OWL Manchester Syntax

(*obo:FOODON_00002809*, *subClassOf*, *obo:RO_0001000*, some, *obo:FOODON_03411347*)

OWL2Vec*: OWL Ontology Embedding

3. Lexical document – sentences of tokens (words)

Solution #1: Transform from structure document

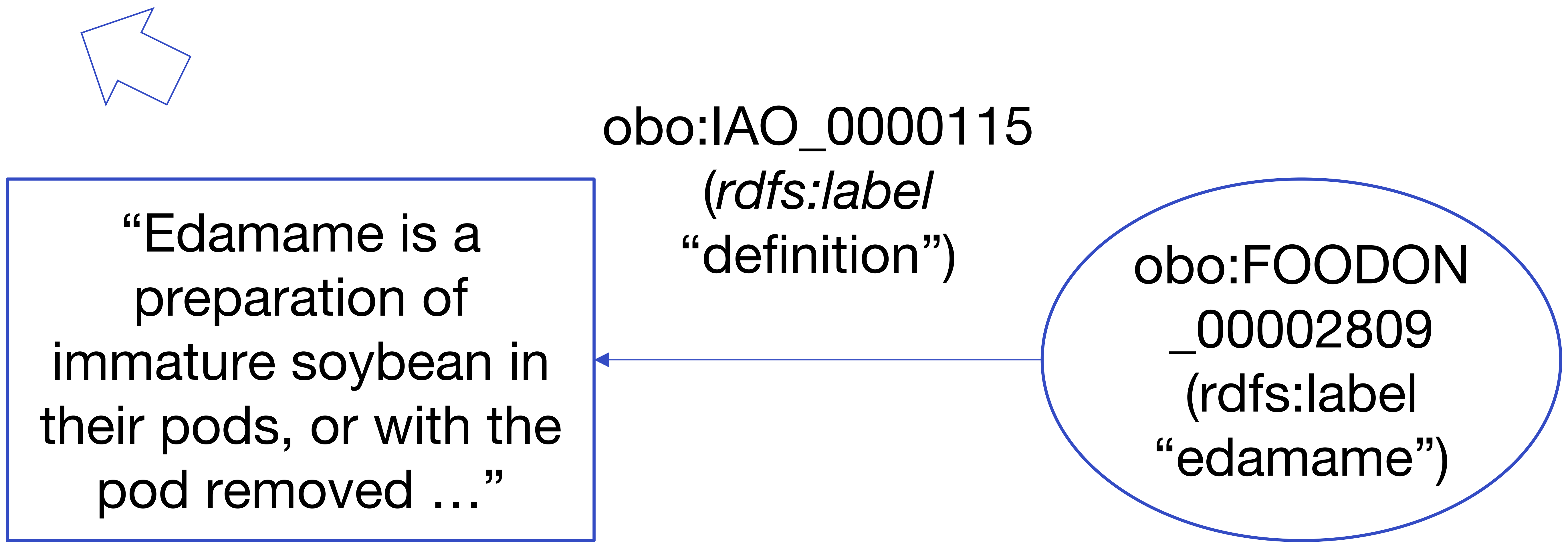
e.g., (*vc:FOOD-4001*, *vc:hasNutrient*, *vc:VitaminC_100*, *vc:amountNutrient*) → (“blonde”, “beer”, “has”, “nutrient”, “vitamin”, “c”, “amount”, “nutrient”)

OWL2Vec*: OWL Ontology Embedding

3. Lexical document – sentences of tokens (words)

Solution #2: Extraction from text of annotation properties

e.g., (“edamame”, “edamame”, “is”, “a”, “preparation”, “of”, “immature”, “soybean”, “in”, “their”, “pods” ...)



OWL2Vec*: OWL Ontology Embedding

4. Combined document – sentences of tokens and IRIs

Solution: Replace one entity in every entity sequence by its words by random selection or traversal

e.g., (*vc:FOOD-4001*, *vc:hasNutrient*, *vc:VitaminC_100*, *vc:amountNutrient*) → (*vc:FOOD-4001*, “has”, “nutrient”, “vitamin”, “c”, “amount”, “nutrient”)

OWL2Vec*: OWL Ontology Embedding

5. Word embedding model (CBOW)

- Optionally, pre-train by text corpus (e.g., from Wikipedia dump)
- Train by the structure, lexical and combined documents
- Entity vector: IRI vector and/or average word vector

Ontology Completion with OWL2Vec*

- Class membership ($a \in C$) and subsumption ($C \sqsubseteq D$) prediction
  - Similar setting as in Box²EL (rank candidate super classes, MRR, Hits@K)
- The candidates
  - Exclude ancestors except for the ground truth by reasoning
  - Consider neighbors or/and similar classes of the ground truth
- Train a classifier (e.g., Random Forest) from declared subsumptions or memberships for prediction (a score s in $[0,1]$)
  - $f: V_C, V_D \rightarrow s$, or $f: V_C, V_a \rightarrow s$
  - The embedding of an entity V: IRI vector, label's token vector, or their concatenation

MANCHESTER 1824

Ontology Completion with OWL2Vec*

	HeLis			
Method	MRR	Hits@1	Hits@5	Hits@10
Transformer (*label*)	0.657	0.515	0.824	0.897
Transformer (*all text*)	0.599	0.390	0.870	0.912
RDF2Vec	0.345	0.219	0.460	0.655
TransE	0.181	0.09	0.232	0.355
TransR	0.298	0.184	0.391	0.559
DistMult	0.253	0.166	0.304	0.437
Quantum Embeding	0.159	0.132	0.163	0.190
Onto2Vec	0.211	0.108	0.268	0.397
OPA2Vec	0.237	0.146	0.286	0.408
OWL2Vec	0.335	0.215	0.397	0.601
Pre-trained *Word2Vec*	0.899	0.877	0.923	0.933
OWL2Vec*	**0.953**	**0.932**	**0.978**	**0.987**

(a) Membership Prediction

← Membership prediction results on the Healthy Lifestyle Ontology (HeLis)

Ontology Completion with OWL2Vec*

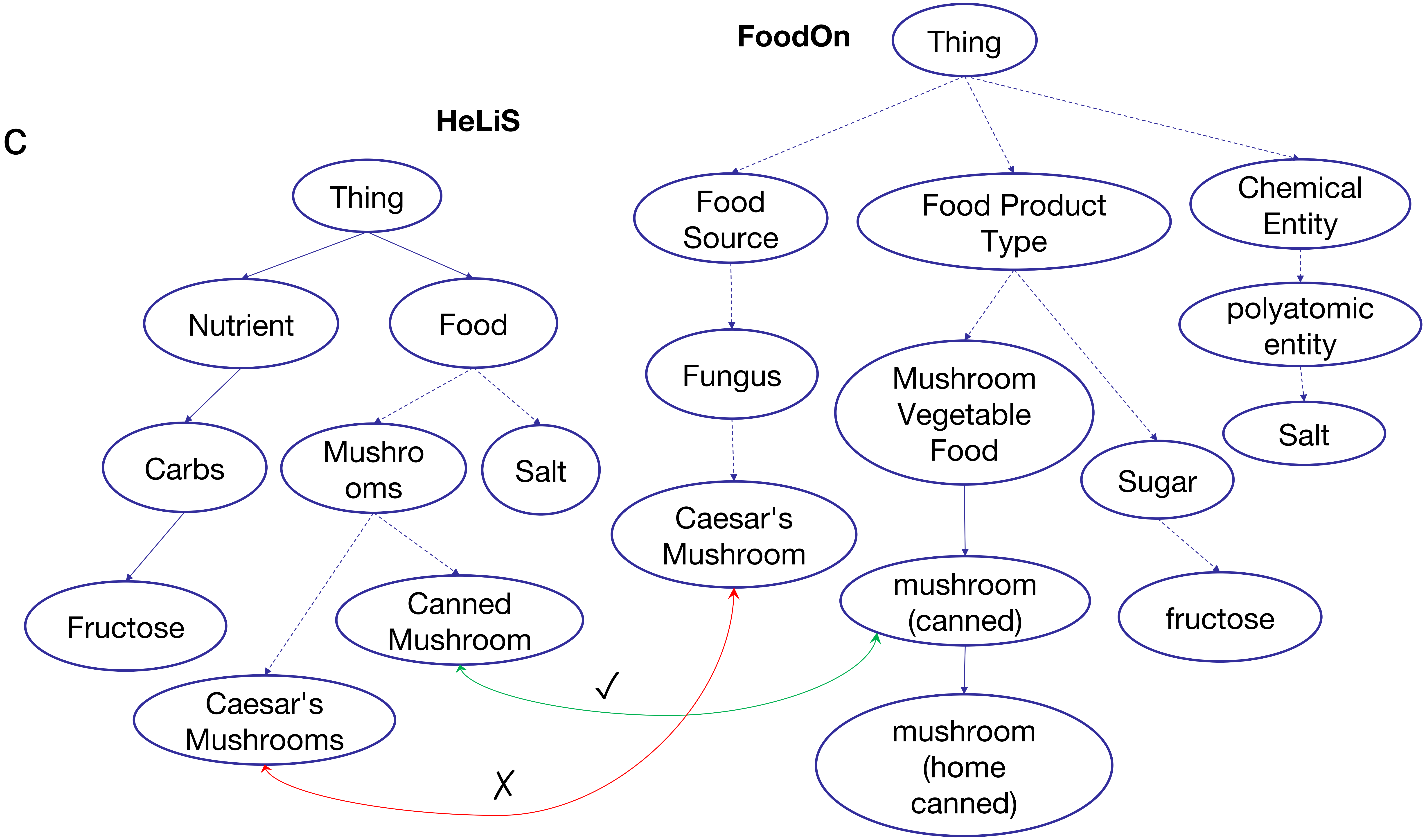
Method	FoodOn			
	MRR	Hits@1	Hits@5	Hits@10
Transformer (*label*)	0.016	0.005	0.027	0.046
Transformer (*all text*)	0.022	0.011	0.032	0.050
RDF2Vec	0.078	0.053	0.097	0.119
TransE	0.029	0.011	0.044	0.065
TransR	0.072	0.044	0.093	0.130
DistMult	0.076	0.045	0.099	0.134
EL Embeding	0.040	0.014	0.067	0.099
Onto2Vec	0.034	0.014	0.047	0.064
OPA2Vec	0.093	0.058	0.117	0.156
OWL2Vec	0.091	0.052	0.121	0.152
Pre-trained *Word-2Vec*	0.136	0.089	0.175	0.227
OWL2Vec*	**0.213**	**0.143**	**0.287**	**0.357**

(b) Subsumption Prediction

← Subsumption prediction results on the Food Ontology (FoodOn)

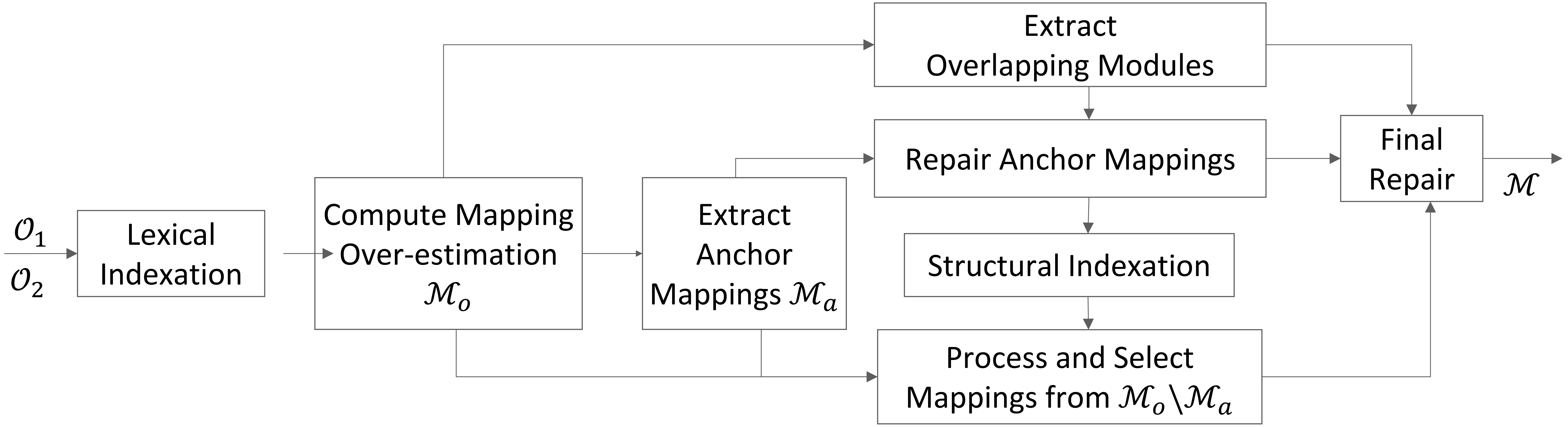
OWL2Vec* for Ontology Alignment

- Find concepts from two ontologies with a specific relationships such as equivalence



OWL2Vec* for Ontology Alignment

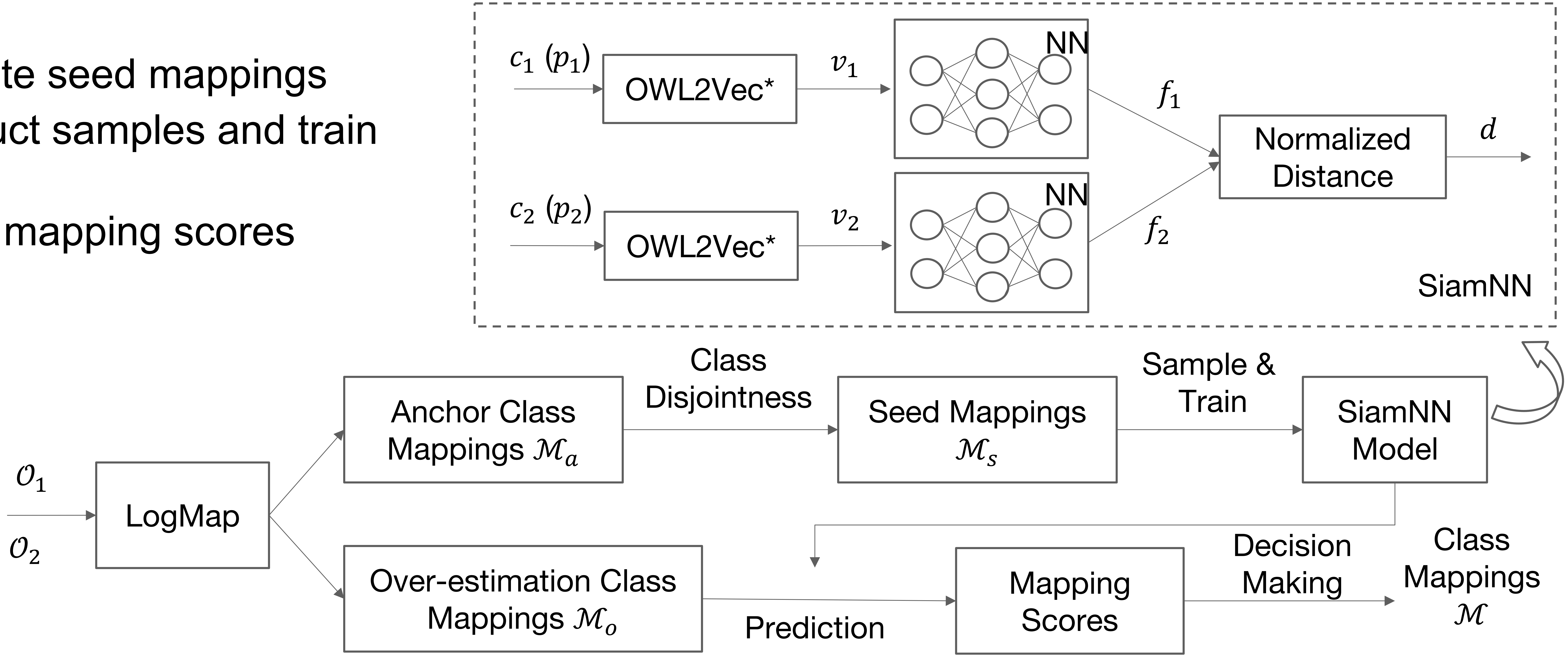
- Traditional system **LogMap**
  - Based on lexical matching and reasoning
  - Over-estimation $\mathcal{M}_o$: high recall, low precision
  - Anchor Mappings $\mathcal{M}_a$: high precision, low recall



Ernesto, J., et al. "Logmap: Logic-based and scalable ontology matching." *ISWC 2011*

LogMap-ML

- Calculate seed mappings
- Construct samples and train model
- Predict mapping scores



Chen, J., et al. "Augmenting ontology alignment by semantic embedding and distant supervision." *ESWC 2021.*

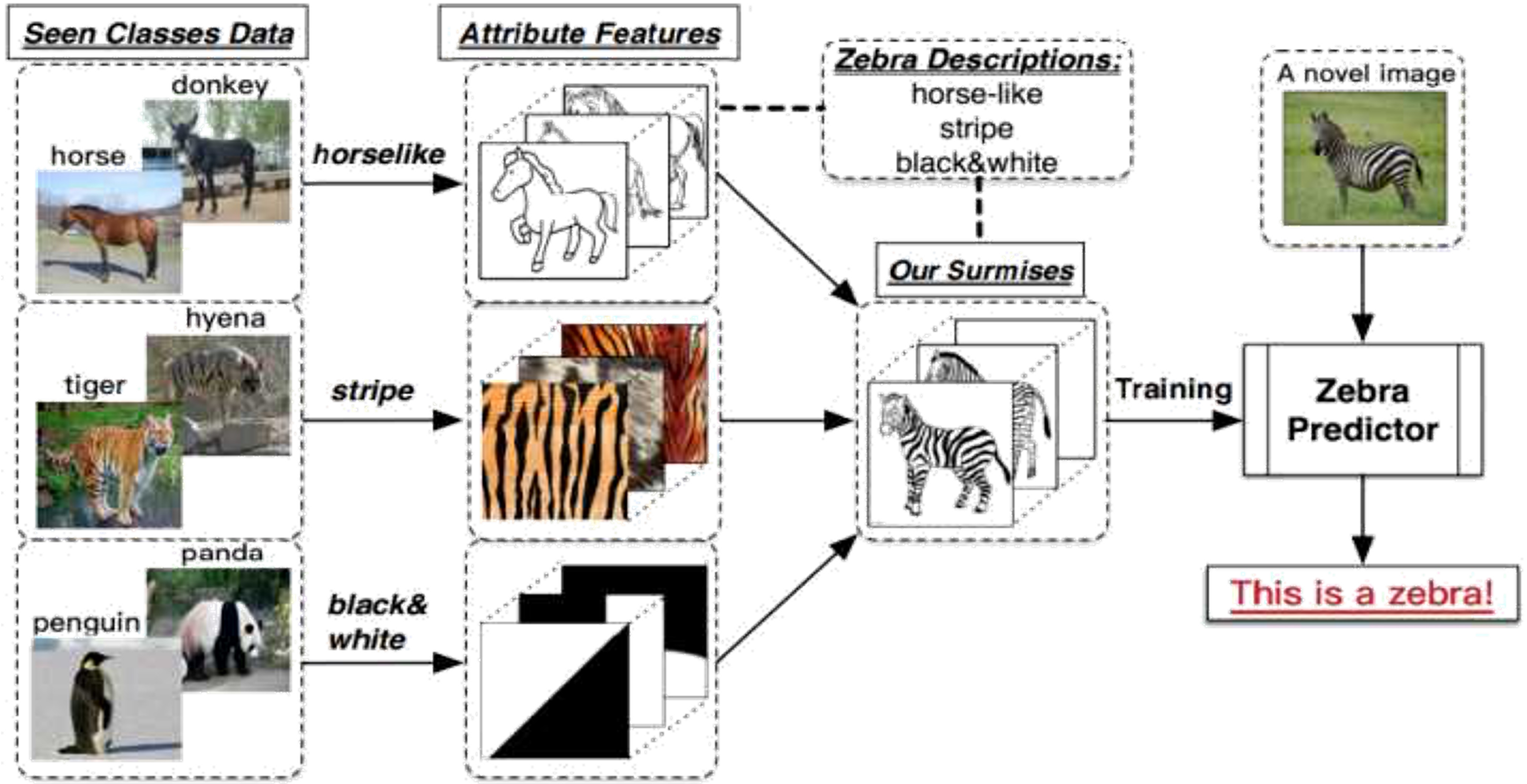
Other Applications of Ontology Embeddings

- Augmenting Machine Learning
  - E.g., injecting external knowledge of classes for zero-shot learning

Chen, J, et al. "Zero-Shot and Few-Shot Learning With Knowledge Graphs: A Comprehensive Survey." Proceedings of the IEEE (2023).

Other Applications of Ontology Embeddings

- What is Zero-shot Learning
  - Predict samples with new classes that have never appeared in training
  - Seen classes vs unseen classes

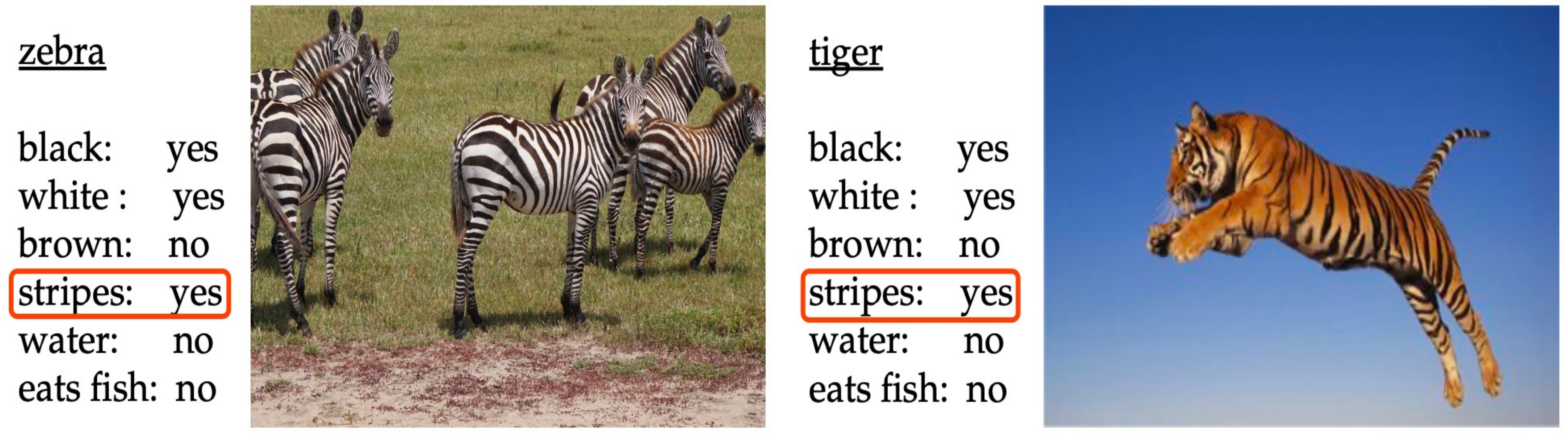


Zero-shot Image Classification

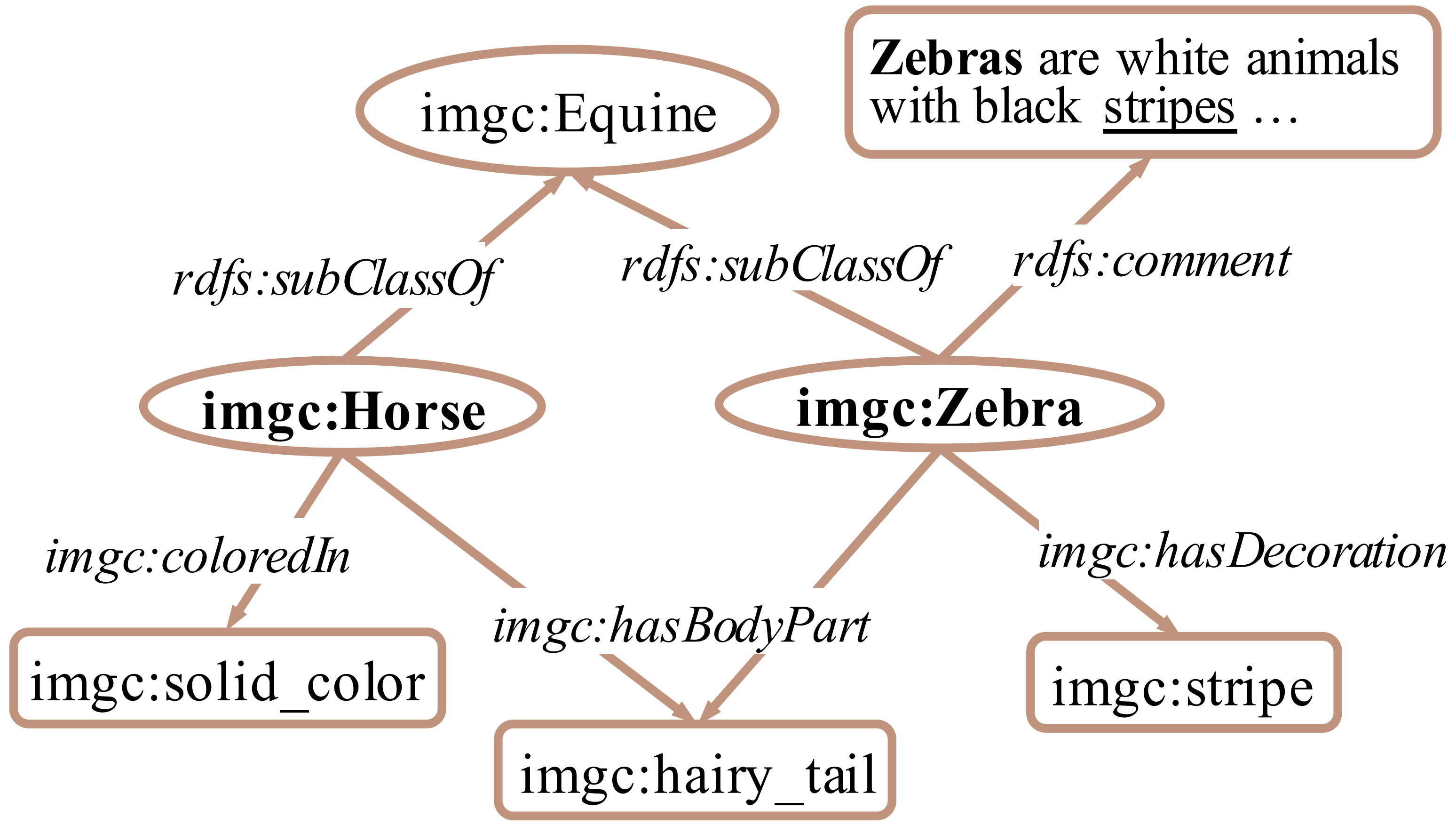
Other Applications of Ontology Embeddings

- **External knowledge** (a.k.a. **side information**) model the relationship between classes, thus enabling the transfer of the model from seen classes to unseen classes.

- Textual description:
  *"**Zebras** are white animals with black stripes, they have larger, rounder ears than horses ..."*



- Attribute descriptions, e.g., visual properties of animals



- Relational Facts, Taxonomy, Literals

"$Zebra \sqsubseteq Equine \sqcap \exists hasTexture.Stripes \sqcap \exists hasHabitat.Meadow$ … "
"$hasUncle \equiv hasParent \circ hasBrother$"

- Logics & rules

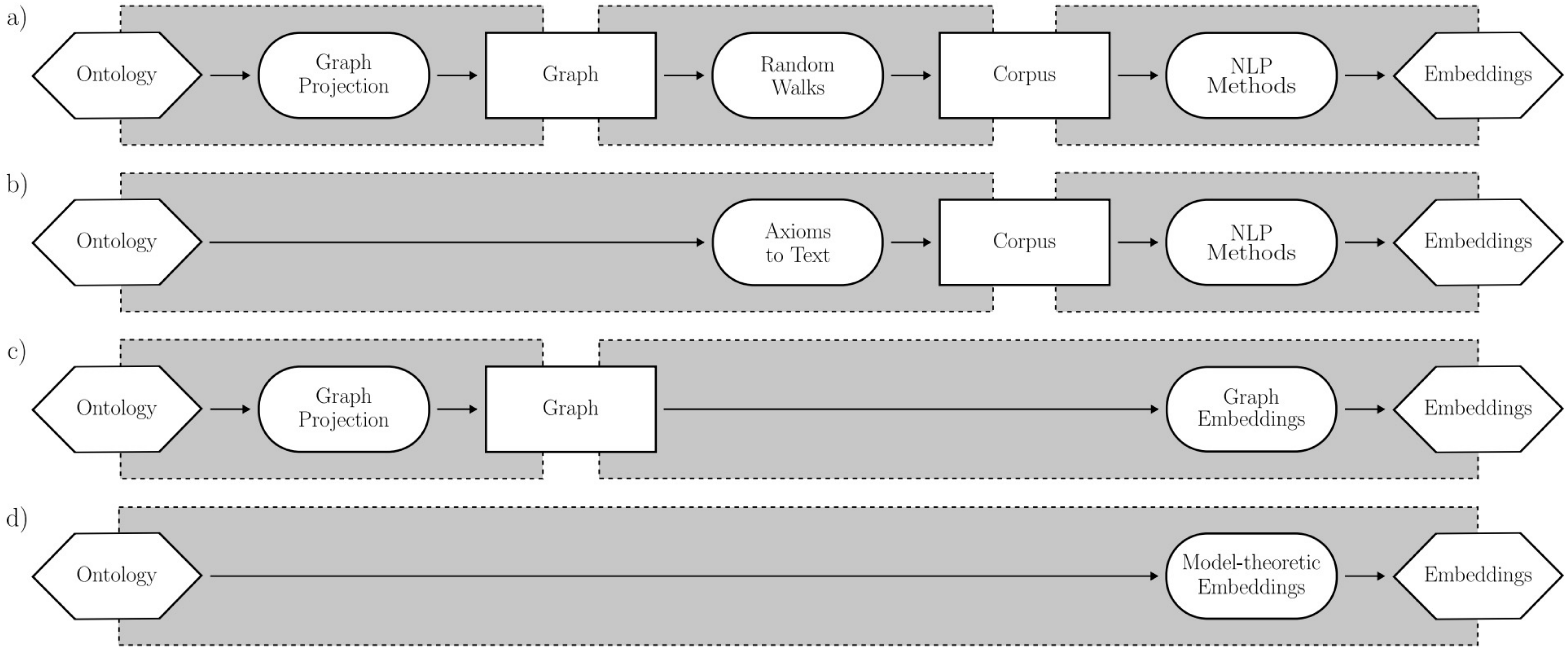
mOWL: A Library for Ontology Embeddings

- Python interfaces for
  - Ontology manipulation and transformation
  - Ontology embedding algorithms and evaluation resources
    - e.g., ELEm, ELBE, BoxEL, Box^{2}EL, OPA2Vec, OWL2Vec*, PPI datasets

https://github.com/bio-ontology-research-group/mowl

Zhapa-Camacho, F., et al. "mOWL: Python library for machine learning with biomedical ontologies." *Bioinformatics* 39.1 (2023): btac811.

mOWL: A Library for Ontology Embeddings



mOWL's workflows for ontology embedding implementation

Summary

- Ontology vs Knowledge Graph
- Geometric modeling
  - Concept as ball: ELEm
  - Concept as box: Box^2EL
  - Evaluation & application (ontology completion)
- Literal-aware ontology embedding
  - OWL2Vec*
  - Evaluation & application (ontology completion & alignment)
- Ontology embedding for zero-shot learning
- mOWL: a library for ontology embedding

The End of Day 4