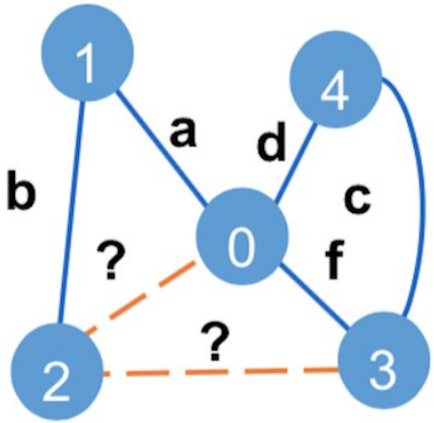
Session 3: Conditional Constraints for KG Embeddings

Michael Weyns, Pieter Bonte, Bram Steenwinckel, Filip De Turck, and Femke Ongenae

Context

KG completion → link prediction

True and false facts required → negative sampling



$$node\ 0 = \begin{bmatrix} 0.1 \\ ... \\ 0.13 \end{bmatrix}$$

$$edge\ a = \begin{bmatrix} 0.32 \\ ... \\ 0.3 \end{bmatrix}$$

SOTA & objectives

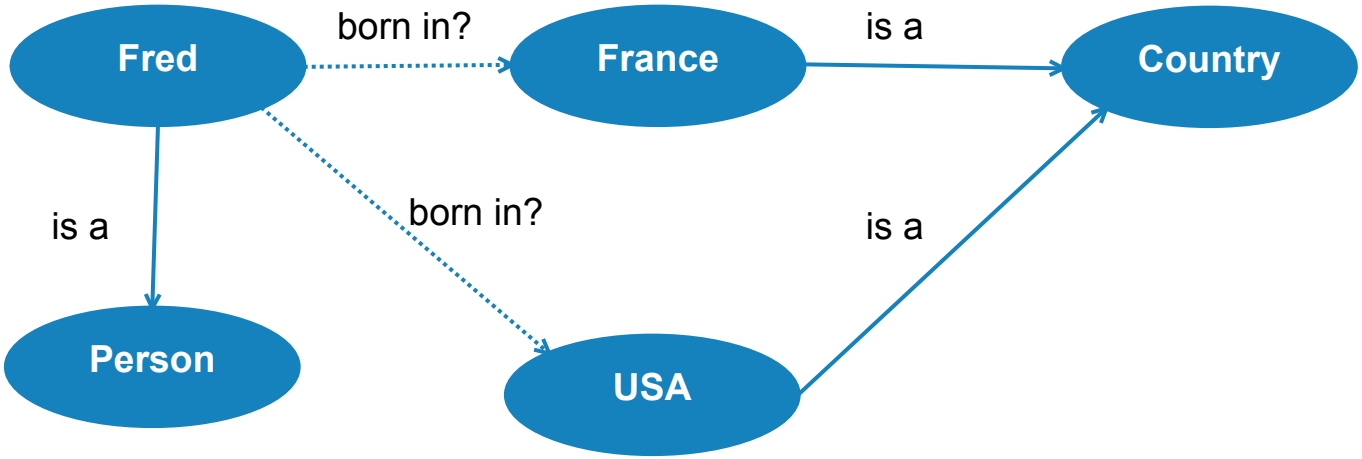
- Exploiting schema to improve negative sampling
- Context-free constraints (RDFS domain and range axioms)
- Closed-world constraints

Objectives:

- Conditional constraints (OWL restrictions)
- Open-world constraints

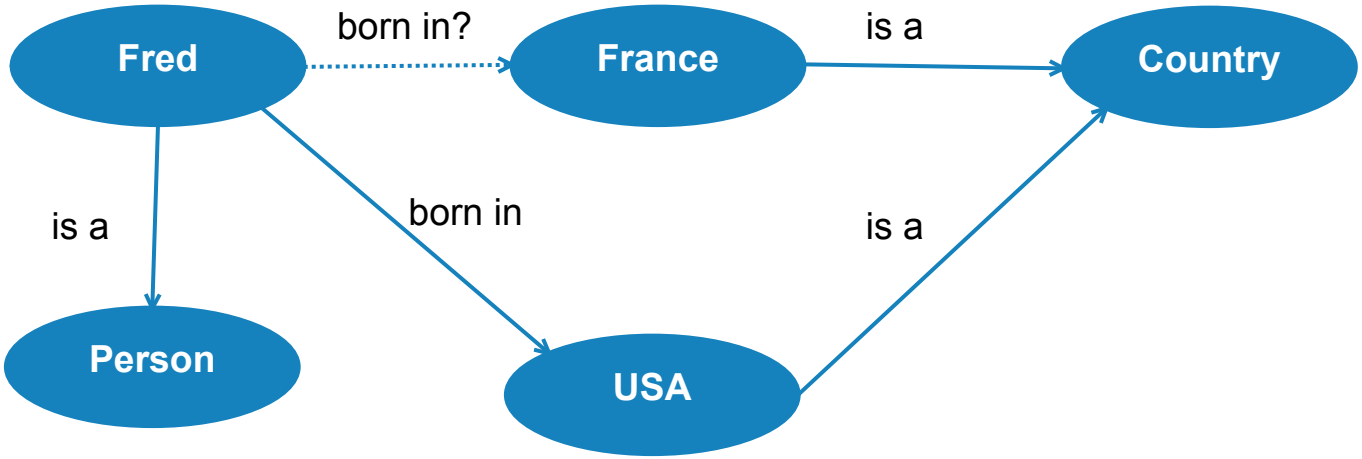
Open World Assumption (OWA)

Incomplete knowledge



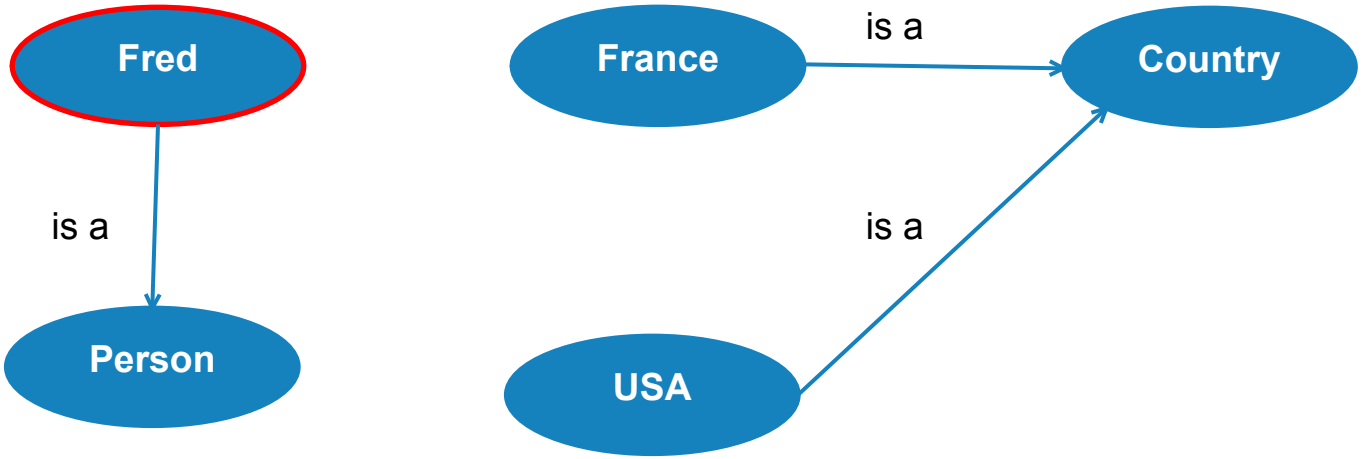
Open World Assumption (OWA)

Monotonicity



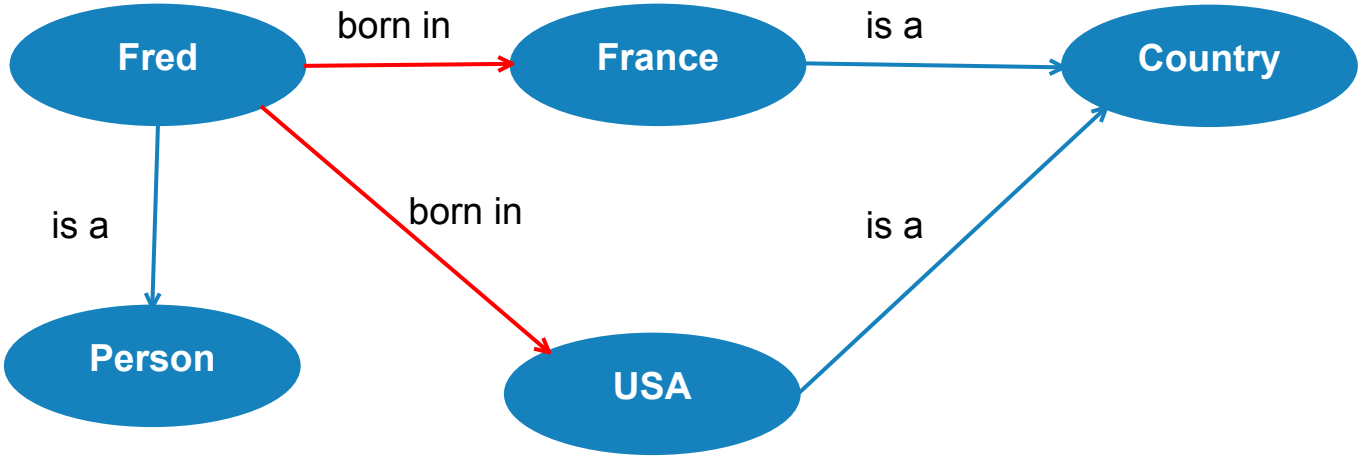
Open World Assumption (OWA) - limits

Inconsistency - restriction on Fred (Person is born in some Country)



Open World Assumption (OWA) - limits

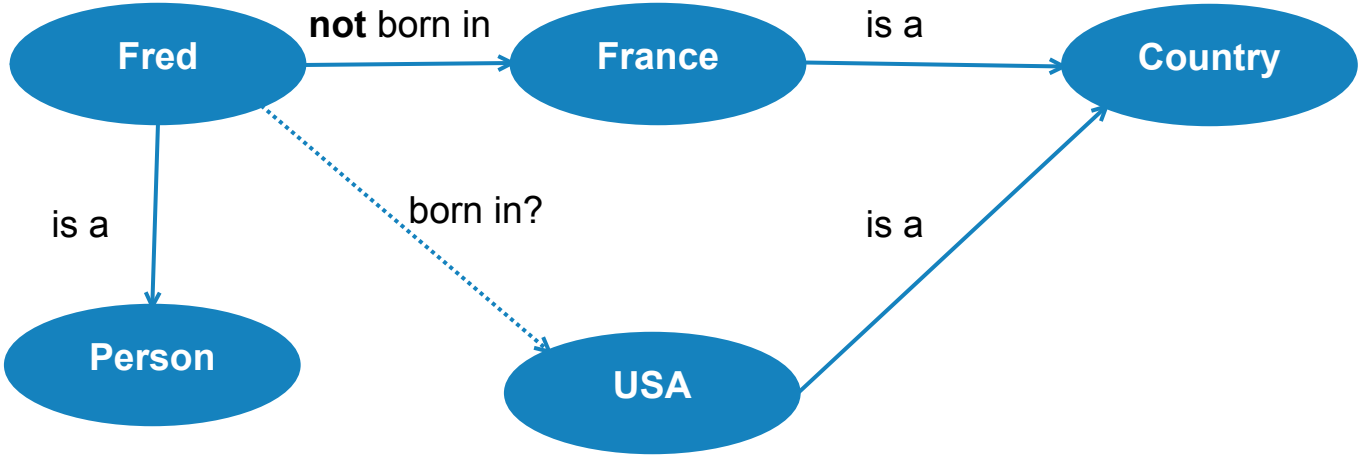
Inconsistency - restriction on Fred (Person is born in max 1 Country)



Open World Assumption (OWA) - limits

Negative property assertions

e.g. NegativeObjectPropertyAssertion(:born_in :Fred :USA)



Negative sampling - SOTA

$$\forall i,k,j,(e_i,r_k,e_j) \in O \implies e_i \notin \mathcal{C}, e_j \notin \mathcal{C}$$
$$\forall (e_i,r_k,e_j) \in O, (e_i,r_k,e_j) \in O^+ \iff y_{ikj} = 1$$
$$\forall (e_i,r_k,e_j) \in O, (e_i,r_k,e_j) \in O^- \implies y_{ikj} = 0$$

$$y_{ikj} = \begin{cases} 1, & \text{if } x_{ikj} \text{ exists} \\ 0, & \text{otherwise} \end{cases}$$

Negative sampling - SOTA

CWA

<Fred, born_in, USA>

<Fred, born_in, France>

<Fred, born_in, Belgium>

<Fred, born_in, England>

<Lucy, born_in, Scotland>

...

Negative sampling - SOTA

Perturbation (+ filtering)

<Fred, born_in, USA>

<**Lucy**, born_in, USA>

<Fred, born_in, **France**>

Negative sampling - SOTA

Bernoulli trick

per relationship *r:*

$$hpt = avg\ \#\ head\ entities/tail\ entity$$
$$tph = avg\ \#\ tail\ entities/head\ entity$$

perturb ***head*** with $prob = \frac{tph}{(tph + hpt)}$

perturb ***tail*** with $prob = \frac{hpt}{(hpt + tph)}$

RDFS axioms

$$\mathcal{K} = \{1 \dots N_r\}, \mathcal{I} = \{1 \dots N_e\}$$

$$\forall k \in \mathcal{K}, \forall c \in \mathcal{C}, \forall (r_k, \text{rdfs:domain}, c) \in \text{TBox} \implies$$
$$\forall i, j \in \mathcal{I}, (e_i, r_k, e_j) \implies (e_i, \text{a}, c)$$
$$\forall k \in \mathcal{K}, \forall c \in \mathcal{C}, \forall (r_k, \text{rdfs:range}, c) \in \text{TBox} \implies$$
$$\forall i, j \in \mathcal{I}, (e_i, r_k, e_j) \implies (e_j, \text{a}, c)$$

OWL axioms

$$\mathcal{K} = \{1 \dots N_r\}, \mathcal{I} = \{1 \dots N_e\}$$

$$\begin{aligned}
&\forall k \in \mathcal{K}, \forall c, c' \in \mathcal{C}, \forall (b(c, r_k), \text{owl:onProperty}, r_k) \in \text{TBox} \ \& \\
&\qquad \forall (b(c, r_k), \text{owl:allValuesFrom}, c') \in \text{TBox} \\
&\qquad \implies \forall i, j \in \mathcal{I}, (e_i, a, c) \implies (e_i, r_k, e_j) \implies (e_j, \text{a}, c') \\
&\forall k \in \mathcal{K}, \forall c, c' \in \mathcal{C}, \forall (b(c, r_k), \text{owl:onProperty}, r_k) \in \text{TBox} \ \& \\
&\qquad \forall (b(c, r_k), \text{owl:someValuesFrom}, c') \in \text{TBox} \\
&\implies \forall i \in \mathcal{I}, \exists j \in \mathcal{I}, (e_i, a, c) \implies (e_i, r_k, e_j) \ \& \ (e_j, \text{a}, c')
\end{aligned}$$

Context-free constraints - SOTA

rdfs:domain is an instance of *rdf:Property* that is used to state that any resource that has a given property *must be* an instance of one or more classes. *rdfs:range* is an instance of *rdf:Property* that is used to state that the values of a property *must be* instances of one or more classes.

$$\forall k \in \mathcal{K}, \forall c \in \mathcal{C}, \forall (r_k, \text{rdfs:domain}, c) \in \text{TBox} \implies$$
$$\forall i, j \in \mathcal{I}, (e_i, \text{a}, c) \implies (e_i, r_k, e_j) \text{ is valid}$$
$$\forall k \in \mathcal{K}, \forall c \in \mathcal{C}, \forall (r_k, \text{rdfs:range}, c) \in \text{TBox} \implies$$
$$\forall i, j \in \mathcal{I}, (e_j, \text{a}, c) \implies (e_i, r_k, e_j) \text{ is valid}$$

Conditional constraints

The *owl:allValuesFrom* restriction requires that for every instance of the class that has instances of the specified property, the values of the property *must all be* members of the class indicated by the *owl:allValuesFrom* clause

The *owl:someValuesFrom* restriction describes a class of all individuals for which at least one value of the property concerned *must be* an instance of the class description or a data value in the data range

$$\begin{aligned}
\forall k \in \mathcal{K}, \forall c, c' \in \mathcal{C}, \forall (b(c, r_k), \text{owl:onProperty}, r_k) \in \text{TBox} \ \& \\
\forall (b(c, r_k), \text{owl:allValuesFrom}, c') \in \text{TBox} \\
\implies \forall i, j \in \mathcal{I}, (e_i, a, c) \implies (e_j, \text{a}, c') \\
\implies (e_i, r_k, e_j) \text{ is valid} \\
\forall k \in \mathcal{K}, \forall c, c' \in \mathcal{C}, \forall (b(c, r_k), \text{owl:onProperty}, r_k) \in \text{TBox} \ \& \\
\forall (b(c, r_k), \text{owl:someValuesFrom}, c') \in \text{TBox} \\
\implies \forall i \in \mathcal{I}, \exists j \in \mathcal{I}, (e_i, r_k, e_j) \ \& \ (e_j, \text{a}, c') \implies (e_i, a, c) \\
\implies (e_i, r_k, e_j) \text{ is valid}
\end{aligned}$$

Constraint-based negative sampling

1. Type inference based on axioms
2. Impose restrictive interpretation
3. Constraint-based negative sampling

≡

Axiomatic consistency checking during perturbation

Constraints - CWA interpretation (SOTA)

$$\forall i, k, j, (e_i, r_k, e_j) \in O \implies e_i \notin \mathcal{C}, e_j \notin \mathcal{C}$$

$$\forall (e_i, r_k, e_j) \in O, (e_i, r_k, e_j) \in O^+ \iff y_{ikj} = 1$$

$$\forall (e_i, r_k, e_j) \in O, (e_i, r_k, e_j) \in \boxed{O^-} \implies y_{ikj} = 0$$

(e_i, r_k, e_j) *is* ***valid***

Constraints - OWA interpretation

$$\forall i,k,j,(e_i,r_k,e_j) \in O \implies e_i \notin \mathcal{C}, e_j \notin \mathcal{C}$$

$$\forall (e_i,r_k,e_j) \in O, (e_i,r_k,e_j) \in O^+ \iff y_{ikj} = 1$$

$$\forall (e_i,r_k,e_j) \in O, (e_i,r_k,e_j) \in O^- \implies y_{ikj} = 0$$

(e_i, r_k, e_j) *is* ***invalid***

Evaluation - datasets

TransE embedding technique

AIFB: research staff, research groups, affiliations, publications

train	**19916 entities**
valid	**2213 entities**
test	**2459 entities**
OWL constraints	**152**

MUTAG: potentially carcinogenic molecules

train	**41999 entities**
valid	**4667 entities**
test	**5185 entities**
RDFS constraints	**5087**

Evaluation - results

Setting	neg ratio	MR	MRR	Hits@10	FN
no constraints	1	667.860	0.230	0.452	222
no constraints	5	**653.403**	**0.340**	**0.572**	647
open world constraints	1	663.658	0.208	0.412	173
open world constraints	5	N/A	N/A	N/A	N/A
closed world constraints	1	801.260	0.265	0.486	468
closed world constraints	5	870.899	0.250	0.468	985

Table 1: Results for AIFB

Evaluation - results

Setting	neg ratio	MR	MRR	Hits@10	FN
no constraints	1	3408.810	0.043	0.089	29
no constraints	5	2341.297	0.093	0.181	167
open world constraints	1	3166.192	0.051	0.105	12
open world constraints	5	2170.180	0.100	0.180	82
closed world constraints	1	3232.653	0.043	0.091	30
closed world constraints	5	**1975.455**	**0.110**	**0.215**	153

Table 2: Results for MUTAG

Conclusions

- AIFB (conditional constraints)
  - OWA interpretation
    - No improvements
    - Decrease in false negatives
  - CWA interpretation
    - Few false negatives: clear improvements
    - Many false negatives: fewer improvements
  - Best setting: no constraints, with high neg ratio

  - Few conditional constraints:
    - Many false negatives (CWA)
    - High computational complexity (OWA)

Conclusions

- MUTAG (context-free constraints)
  - OWA interpretation
    - Clear improvements
    - Decrease in false negatives
  - CWA interpretation
    - Clear improvements
    - No increase in false negatives
  - Best setting: CWA constraints, with high neg ratio

  - Sufficient conditional constraints:
    - Consistent number of false negatives (CWA)
    - Consistent computational complexity (OWA)

Future work

- Context-free ↔ conditional constraints (same dataset comparison)
- Rejection hyperparameter
- Effects on other embedding strategies

Discussion

Thank you very much for listening.
Any questions?

embracing a better life