Abstract

The datasets provided by the Linked Data community currently form the world's largest, freely available, decentralised and interlinked knowledge bases. However, to be able to benefit from this knowledge in a specific use-case, one typically needs to understand the modelling of the knowledge and formulate appropriate SPARQL queries.

In order to ease this process, we developed an evolutionary algorithm that learns such SPARQL queries (graph patterns) for pairwise relations between source and target entities. Given a training list of source-target-pairs, our algorithm learns a predictive model, which given a new source entity predicts target entities analogously to the training examples.

In our demo paper we present a high level overview over our graph pattern learner and show its application to simulate human associations (e.g., "fish - water"). In the demo users can choose a semantic entity (e.g., dbr:Fish) as stimulus and let the learned model predict human-like responses (e.g., dbr:Water).

Predicting Human Associations with Graph Patterns Learned from Linked Data

Jörn Hees, Rouven Bauer, Joachim Folz, Damian Borth, Andreas Dengel

Try it out → https://w3id.org/associations/#demos ← Try it out

Demo 46

Introduction:

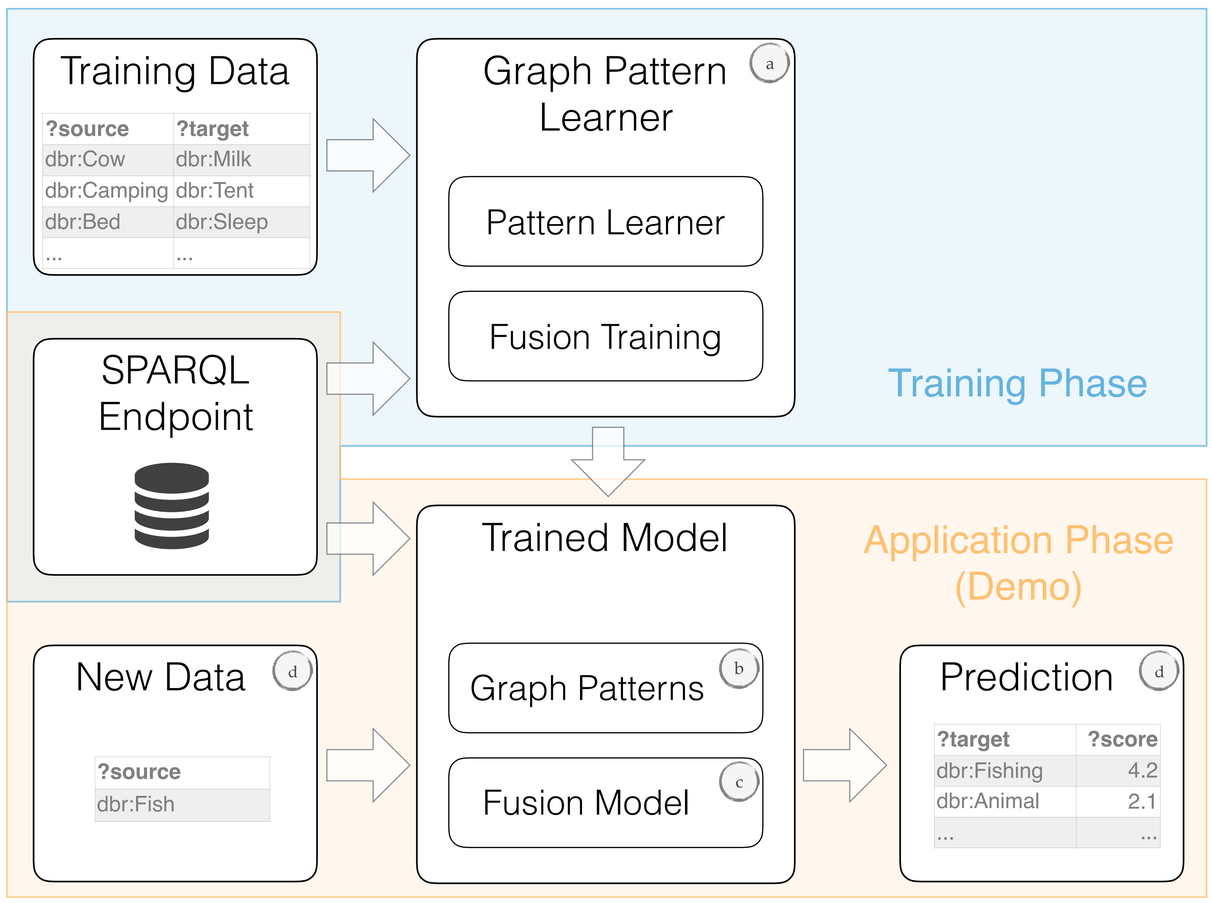
Task:

- Simulate human associations
- For a given source node predict target nodes as humans would

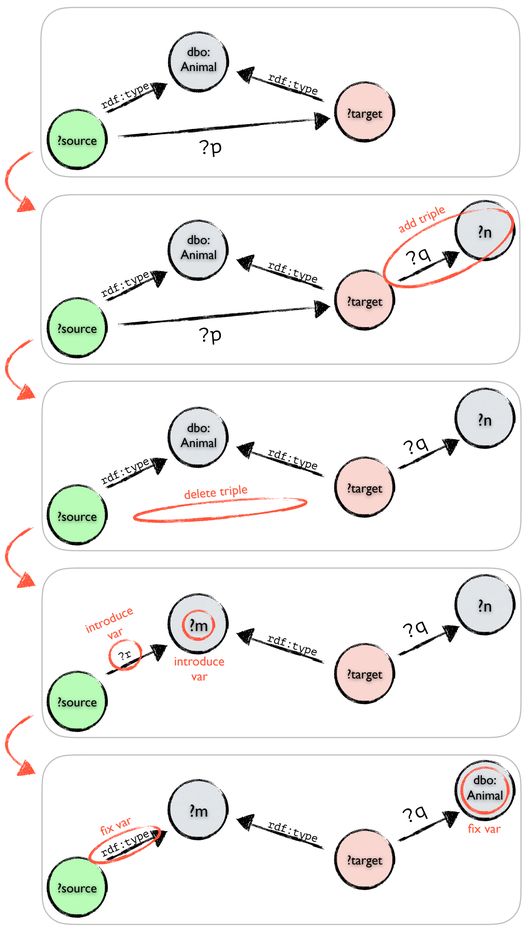
Approach:

- Training Data:
  - Node Pairs (associated)
  - SPARQL Endpoint
- Graph Pattern Learner (a)
  - Learns SPARQL Queries (b)
  - In ensemble can predict training target for corresponding training source (c)
- Apply trained model to user inputs (d)

Machine Learning Outline:



Graph Pattern Learner Mutations: (e)

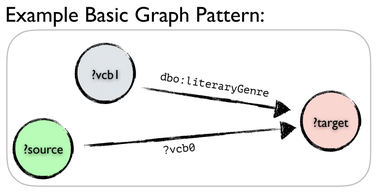


Graph Pattern Learner: (a)

- Evolutionary Algorithm
  - Individuals: SPARQL BGPs
  - Fitness: "Good Patterns"
  - Mutations (e)
  - Mating: Combining Triples (randomly)
  - Init: Random length paths
- Coverage by restarting training runs with remaining pairs
- Clustering of learned patterns via training pair fulfillment (ASK queries)
- Fusion training on target candidate vectors wrt. learned patterns

Individuals: SPARQL BGPs (b)

Example Basic Graph Pattern:




```
{
   ?source ?vcb0 ?target .
   ?vcb1 dbo:literaryGenre ?target .
}
```

Pattern during Training:

```
ASK {
   dbr:Cow ?vcb0 dbr:Milk .
   ?vcb1 dbo:literaryGenre dbr:Milk.
}
```

```
SELECT (COUNT(?target) as ?c) {
   dbr:Cow ?vcb0 ?target .
   ?vcb1 dbo:literaryGenre ?target .
}
```

Pattern used for prediction: (d)

```
SELECT ?target WHERE {
   dbr:Fish ?vcb0 ?target .
   ?vcb1 dbo:literaryGenre ?target .
}
```

Fitness: What is a "Good Pattern?"

- How often is a response reached? max
- How many other nodes are reached? min
- How many nodes need to be expanded? min
- How long does a query take? min

?source	?target
dbr:Cow	dbr:Milk
dbr:Camping	dbr:Tent
dbr:Expense	dbr:Money
dbr:Bed	dbr:Sleep
dbr:Pupil	dbr:Eye
...	...

Learned Patterns form a Feature Space

- Training pairs and patterns form a boolean vector-space wrt. SPARQL ASK queries
  - We can use this to cluster similar patterns

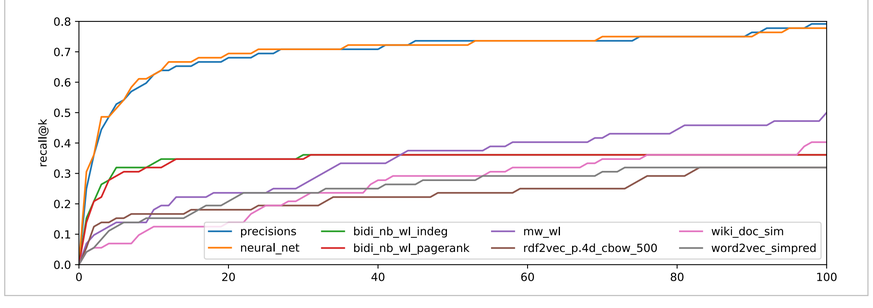
?source	?target	P1	P2	P3	P4	...	P100
dbr:Cow	dbr:Milk	1	0	1	0	...	0
dbr:Camping	dbr:Tent	0	1	1	1	...	0
dbr:Expense	dbr:Money	1	0	0	0	...	1
dbr:Bed	dbr:Sleep	0	1	1	1	...	0
dbr:Pupil	dbr:Eye	1	1	1	1	...	1
...	...	...	...	...	...	...	...

- Dual function of patterns in the vector space:
  - Dimensions & Generating target candidates
  - Used for fusion training with supervised machine learning techniques.

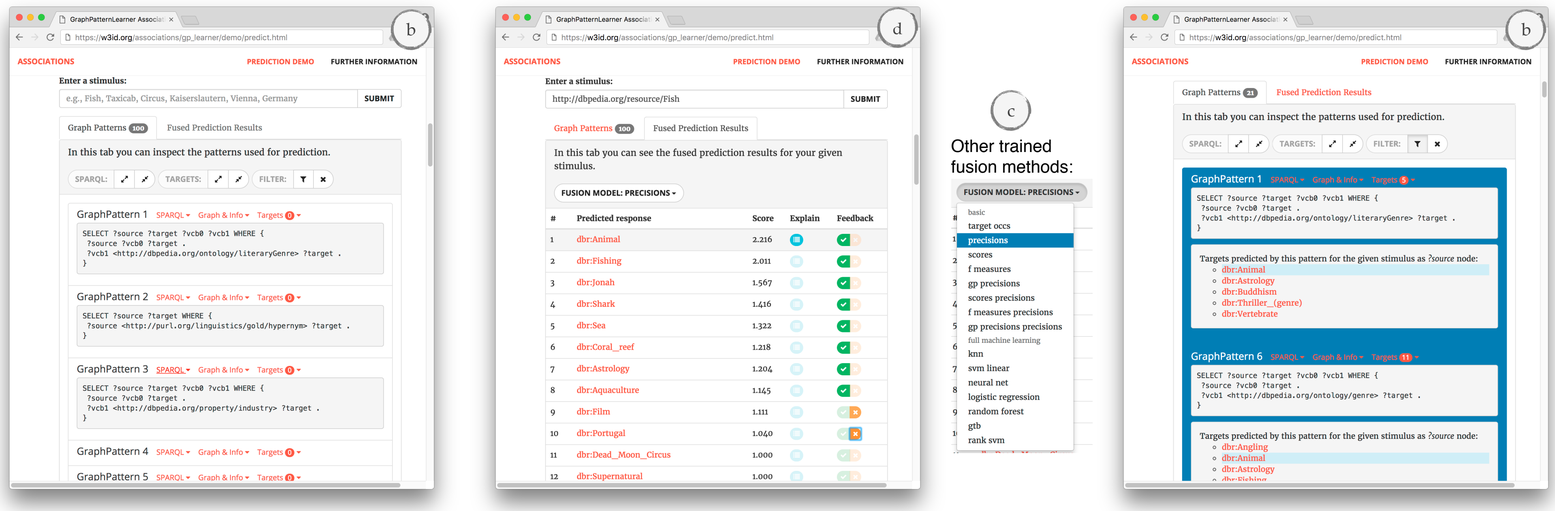
Evaluation:

Avg. Inter-Human Agreement: ~ 32 %

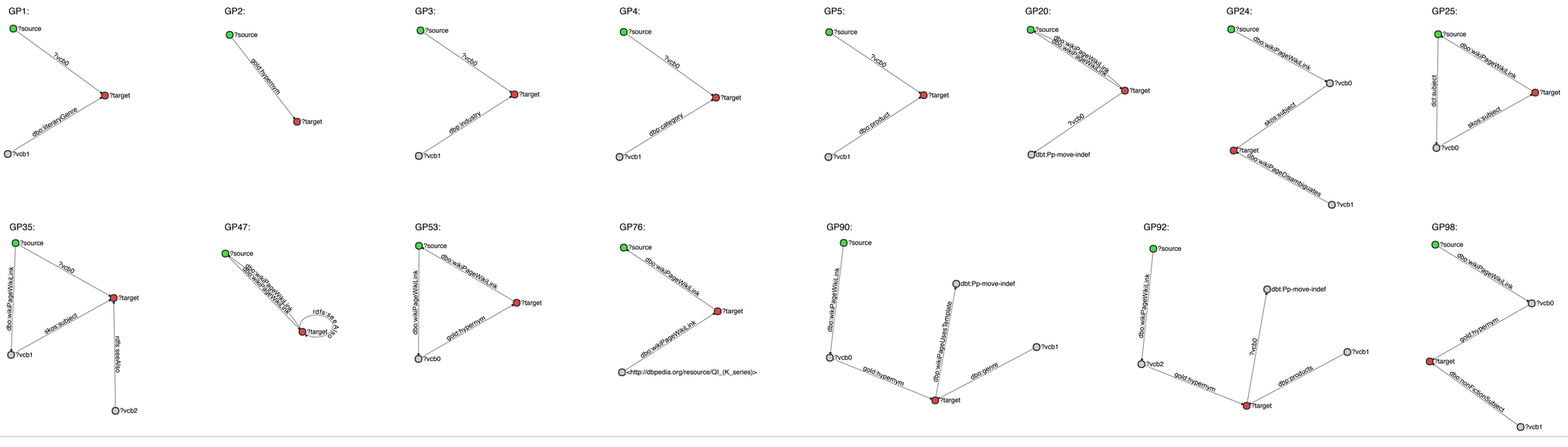
Method	Rec@1	Rec@2	Rec@3	Rec@5	Rec@10	MAP	NDCG
DocSim	4.2%	5.6%	5.6%	6.9%	12.5%	6.6%	12.5%
Word2Vec	4.2%	5.6%	8.3%	12.5%	15.3%	7.8%	12.5%
RDF2Vec	5.6%	12.5%	13.9%	15.3%	16.7%	10.3%	14.4%
MW	6.9%	9.7%	11.1%	13.9%	18.1%	11.0%	17.9%
NB Bidi WL PR	13.9%	20.8%	22.2%	29.2%	31.9%	20.2%	23.8%
NB Bidi WL InDeg	**15.3%**	**20.8%**	**26.4%**	**31.9%**	**33.3%**	**21.4%**	**24.8%**
gpl + precisions	25.0%	36.1%	44.4%	52.8%	62.5%	37.1%	46.0%
gpl + neural net	**30.6%**	**36.1%**	**48.6%**	**51.4%**	**62.5%**	**40.3%**	**48.3%**



Demo in action:



Examples of Learned Graph Patterns in the Trained Model: (b)



More Info & Contact:

w3id.org/associations

joern.hees@dfki.de

German Research Center for Artificial Intelligence

TECHNISCHE UNIVERSITÄT KAISERSLAUTERN