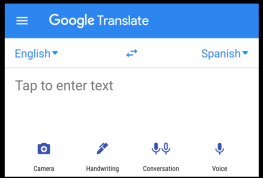
Learning Symbolic Latent Representations for Relational Data

Sebastijan Dumančić
Public PhD Defence
December 11, 2018

Overture

21st CENTURY - AGE OF ARTIFICIAL INTELLIGENCE

AlphaGo



alexa

Powered by machine learning

Systems that improve their performance with experience

Acquiring new skills and knowledge

Machine learning

Systems that improve the performance with experience

Machine learning

Systems that improve the performance with experience

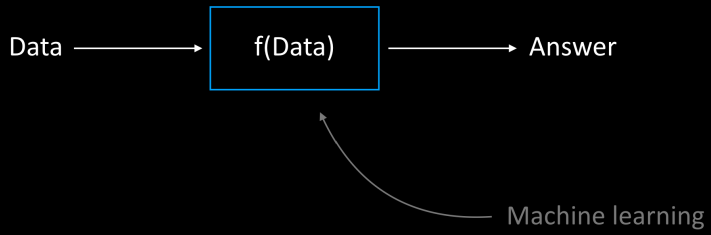
ID	Features						Target
	legs	fluffiness	diet	playful	spits?	bites?	
cat	4	4	carnivore	sometimes	no	sometimes	✓
dog	4	5	carnivore	yes	no	no	✓
duck	2	2	herbivore	no	no	no	×
				...			
alpaca	4	5	herbivore	yes	sometimes	no	?

Systems that improve the performance with experience

ID	Features						Target
	legs	fluffiness	diet	playful	spits?	bites?	
cat	4	4	carnivore	sometimes	no	sometimes	✓
dog	4	5	carnivore	yes	no	no	✓
duck	2	2	herbivore	no	no	no	×
				...			
alpaca	4	5	herbivore	yes	sometimes	no	?

How well can I answer
a question
about the data?

Systems that improve the performance with experience



`IF` fluffy `THEN` good pet

Systems that improve the performance with experience

ID	Features						Target
	legs	fluffiness	diet	playful	spits?	bites?	
cat	4	4	carnivore	sometimes	no	sometimes	✓
dog	4	5	carnivore	yes	no	no	✓
duck	2	2	herbivore	no	no	no	×
				...			
alpaca	4	5	herbivore	yes	sometimes	no	?

How well can I answer
a question
about the data?

Data representation

Why do we care about representations?

Problem representation

Which task is easier to solve?

27 + 15 + 9	CIX + L + IX

Problem representation

Which task is easier to solve?

27 + 15 + 9	CIX + L + IX

Cognitive ease vs cognitive strain [Kahneman,*Thinking fast and slow*]

Problem representation

How we represent a problem matters:

- math problems

$$5 \times 2$$

vs.

$$2 + 2 + 2 + 2 + 2$$

Problem representation

How we represent a problem matters:

- math problems
- searching

Problem representation

How we represent a problem matters:

- math problems
- searching
- communication

a bird of black and white colour, doesn't fly but swims, ...

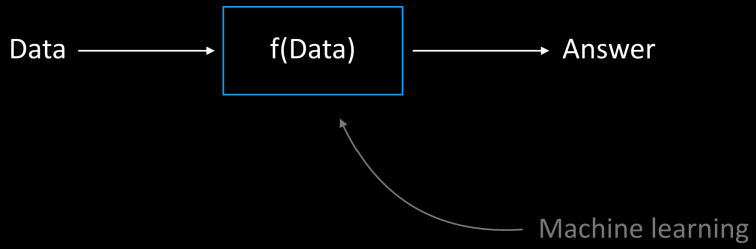
Problem representation

How we represent a problem matters:

- math problems — compact representation
- searching — better performance
- communication — abstraction

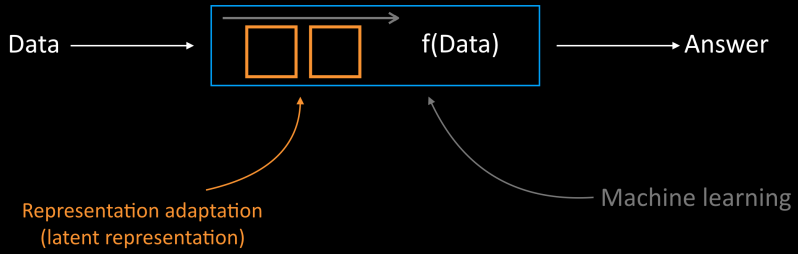
Problem representation

- Representations are important
- but machine learning cannot adapt them



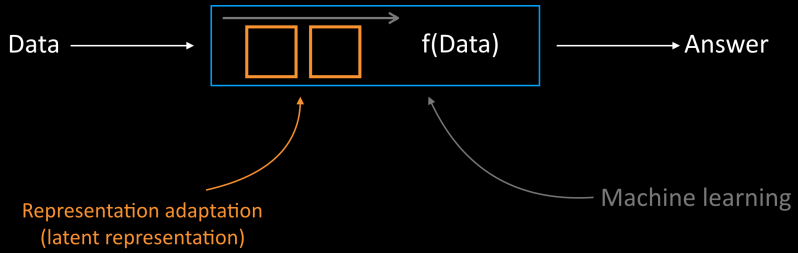
Problem representation

- Representations are important
- but machine learning cannot adapt them
- learning representations with deep learning



Problem representation

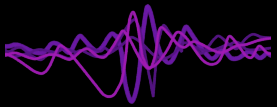
- Representations are important
- but machine learning cannot adapt them
- learning representations with deep learning

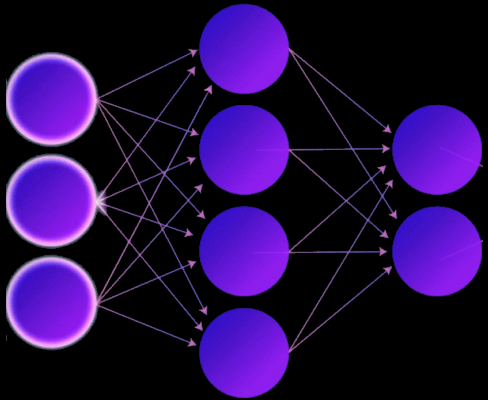


First ingredient of this thesis

Deep learning

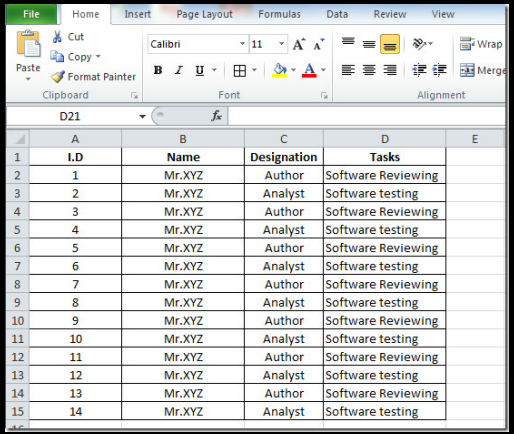
Deep learning: Learning representations from massive sensory data





The shapes of data

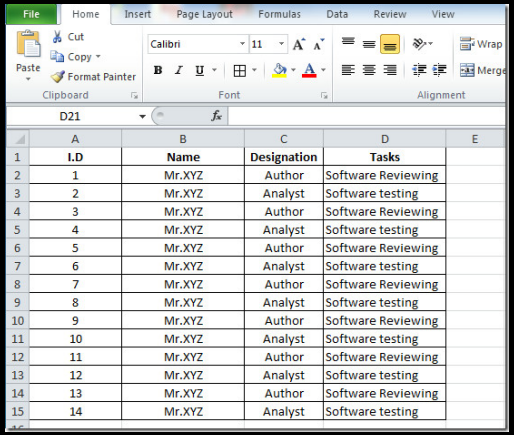
How does data look like?



	A	B	C	D	E
1	I.D	Name	Designation	Tasks	
2	1	Mr.XYZ	Author	Software Reviewing	
3	2	Mr.XYZ	Analyst	Software testing	
4	3	Mr.XYZ	Author	Software Reviewing	
5	4	Mr.XYZ	Analyst	Software testing	
6	5	Mr.XYZ	Author	Software Reviewing	
7	6	Mr.XYZ	Analyst	Software testing	
8	7	Mr.XYZ	Author	Software Reviewing	
9	8	Mr.XYZ	Analyst	Software testing	
10	9	Mr.XYZ	Author	Software Reviewing	
11	10	Mr.XYZ	Analyst	Software testing	
12	11	Mr.XYZ	Author	Software Reviewing	
13	12	Mr.XYZ	Analyst	Software testing	
14	13	Mr.XYZ	Author	Software Reviewing	
15	14	Mr.XYZ	Analyst	Software testing	

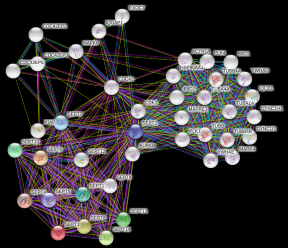
How does data look like?

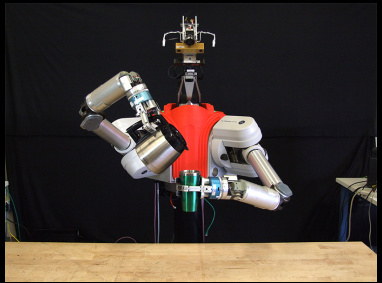
What about more complicated data?



I.D	Name	Designation	Tasks
1	Mr.XYZ	Author	Software Reviewing
2	Mr.XYZ	Analyst	Software testing
3	Mr.XYZ	Author	Software Reviewing
4	Mr.XYZ	Analyst	Software testing
5	Mr.XYZ	Author	Software Reviewing
6	Mr.XYZ	Analyst	Software testing
7	Mr.XYZ	Author	Software Reviewing
8	Mr.XYZ	Analyst	Software testing
9	Mr.XYZ	Author	Software Reviewing
10	Mr.XYZ	Analyst	Software testing
11	Mr.XYZ	Author	Software Reviewing
12	Mr.XYZ	Analyst	Software testing
13	Mr.XYZ	Author	Software Reviewing
14	Mr.XYZ	Analyst	Software testing

Data in the real world

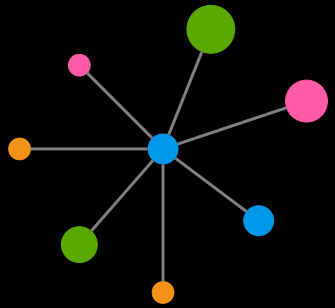




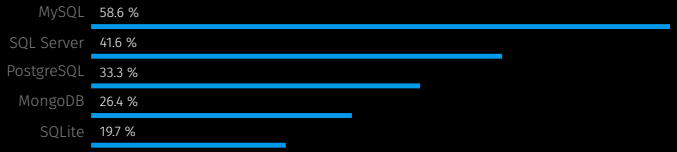
How does real data look like?

Real data contains objects and relationships

- graphs
- networks
- relational databases

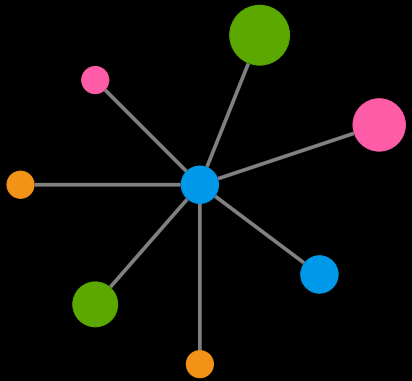


Most commonly used storage technologies (Stackoverflow)



Logic - a powerful way to represent data

Statistical Relational Learning = Logic + Machine learning (+ probability)

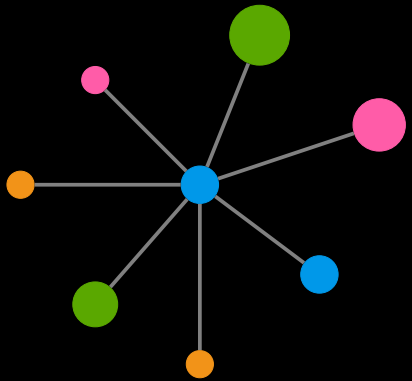


```
likes( ● , ● ).
likes( ● , ● ).
```

Logic - a powerful way to represent data

Logic - a powerful way to represent data

Statistical Relational Learning = Logic + Machine learning (+ probability)



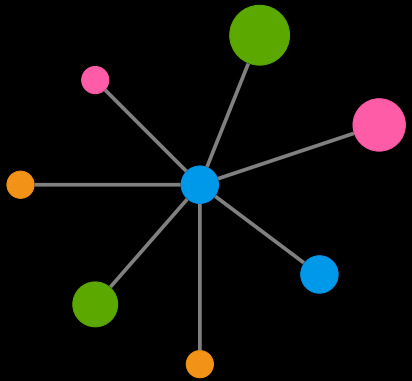
```
likes( ●, ● ).
likes( ●, ● ).

friends(X,Y) :- likes(X,Y).
friends(X,Y) :-
likes(X,Z),friends(Z,Y).
```

Machine Learning: finding predictive formulas

Logic - a powerful way to represent data

Statistical Relational Learning = Logic + Machine learning (+ probability)



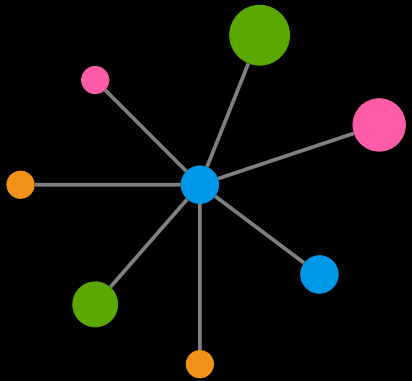
```
likes( ●, ● ).
likes( ●, ● ).

friends(X,Y) :- likes(X,Y).
friends(X,Y) :-
likes(X,Z),friends(Z,Y).
```

Machine Learning: finding predictive formulas

Logic - a powerful way to represent data

Statistical Relational Learning = Logic + Machine learning (+ probability)



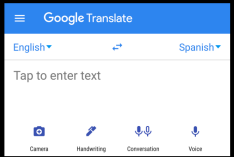
```
likes( ●, ● ).
likes( ●, ● ).

friends(X,Y) :- likes(X,Y).
friends(X,Y) :-
likes(X,Z),friends(Z,Y).
```

Second ingredient of this thesis

WHY DO WE CARE ABOUT LOGIC?

AlphaGo



alexa

WHY DO WE CARE ABOUT LOGIC?

AlphaGo

Machine learning:

- good at pattern matching

Logic:

- good at reasoning

WHY DO WE CARE ABOUT LOGIC?

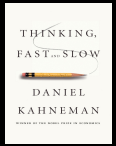
AlphaGo

Machine learning:

- good at pattern matching
- *System 1* – fast but simple

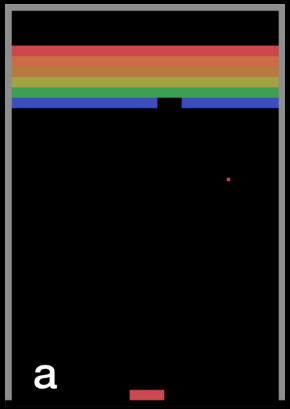
Logic:

- good at reasoning
- *System 2* – slow but complex



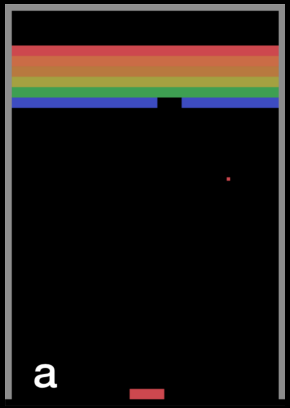
Logic allows to transfer knowledge to different scenarios

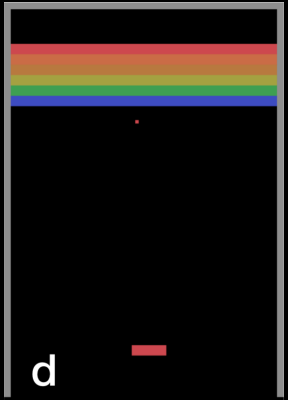
Logic allows to transfer knowledge to different scenarios



vicarious®

Logic allows to transfer knowledge to different scenarios





vicarious®

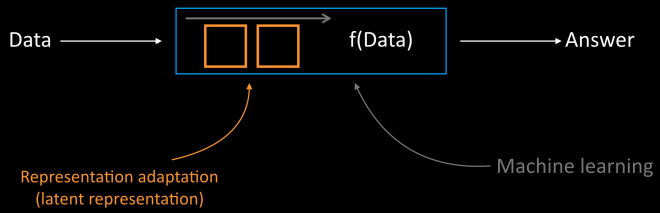
Learning Symbolic Latent Representations for Relational Data

Problem statement

Learning Symbolic Latent Representations for Relational Data

Learning Symbolic Latent Representations for Relational Data

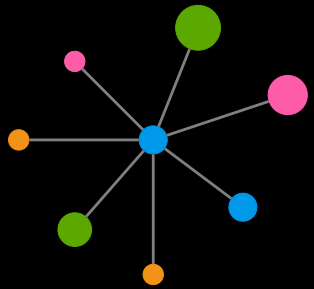
How do we transform the data so that learning is easier?



Learning Symbolic Latent Representations for Relational Data

How do we transform the data so that learning is easier?

When data contains objects and relations

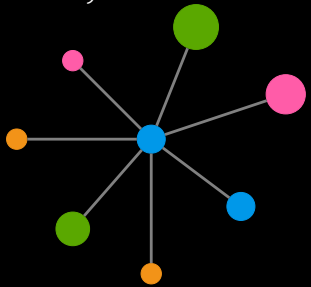


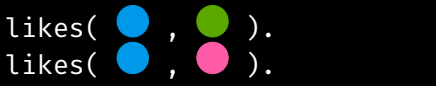
Learning Symbolic Latent Representations for Relational Data

How do we transform the data so that learning is easier?

When data contains objects and relations

in a symbolic way?






```
friends(X,Y) :- likes(X,Y).
friends(X,Y) :-
likes(X,Z),friends(Z,Y).
```

This thesis is the exploration of ideas

PROBLEM How do we learn symbolic latent representation for relational data?

CONTRIBUTION 1 A versatile relational clustering framework
CONTRIBUTION 2 Learning representations through symmetries
CONTRIBUTION 3 Auto-encoding Logic Programs
CONTRIBUTION 4 A comparison of symbolic and sub-symbolic approaches

Contribution 1

Relational clustering framework

Learning representation by clustering

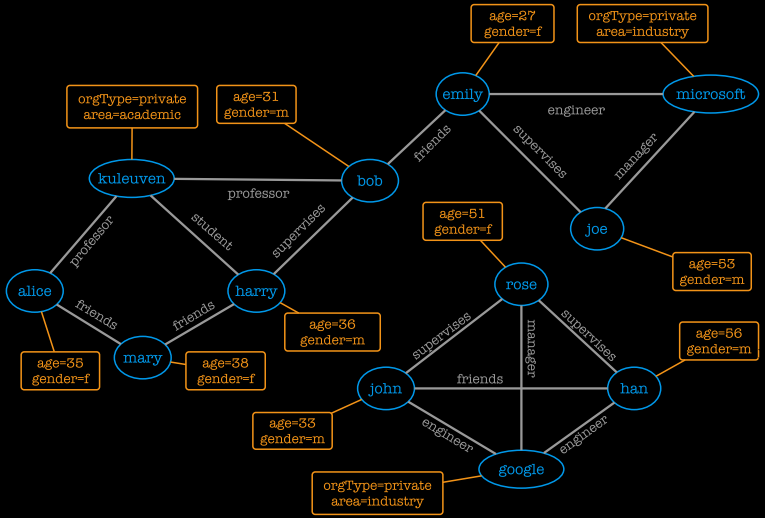
People identify categories by grouping similar objects

Learning representation by clustering

People identify categories by grouping similar objects

Clustering relational data

Groups?

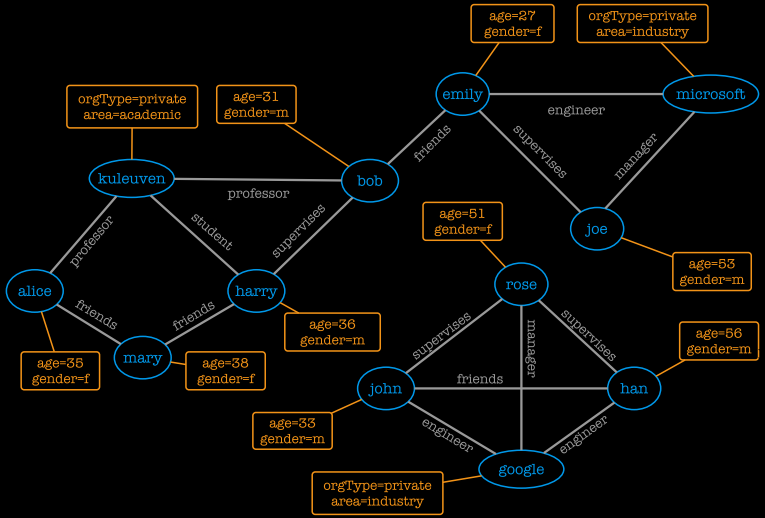


Clustering relational data

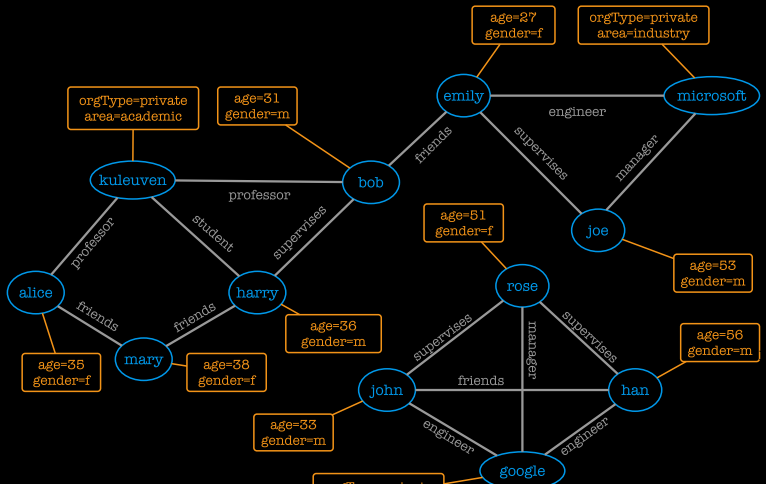
Groups?

Existing methods impose fixed view:

- graph kernels
- relational methods
- graph partitioning

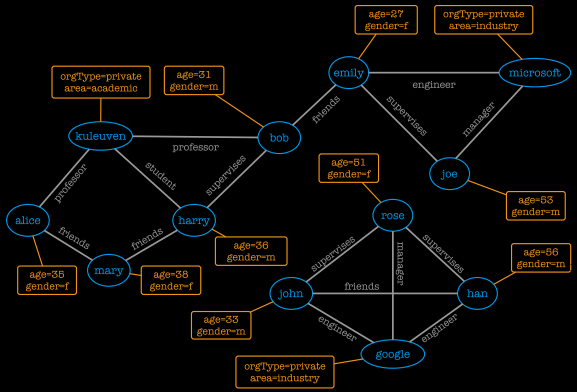


How do we find (various) groups in relational data?



Contribution 1: versatile clustering framework

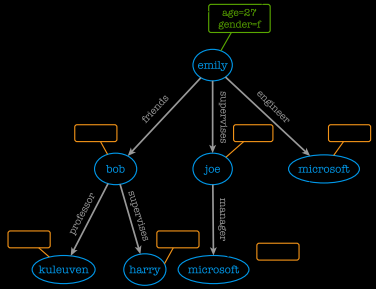
Main idea: disentangle different sources of similarity



Contribution 1: versatile clustering framework

Main idea: disentangle different sources of similarity

STEP 1 build a neighbourhood tree

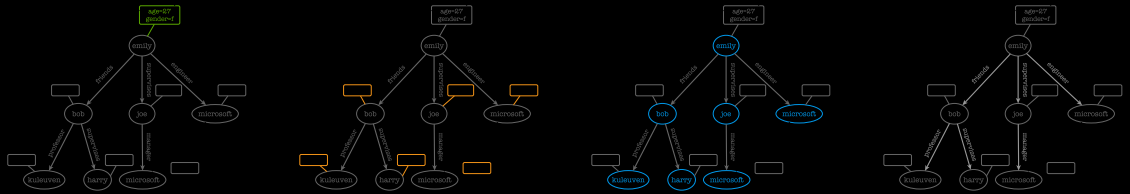


Contribution 1: versatile clustering framework

Main idea: disentangle different sources of similarity

STEP 1 build a neighbourhood tree

STEP 2 split into semantic components

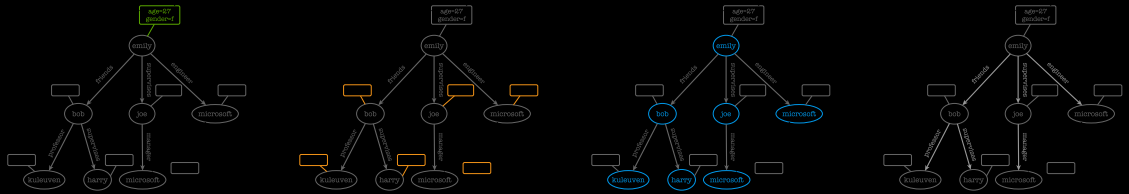


Contribution 1: versatile clustering framework

Main idea: disentangle different sources of similarity

STEP 1 build a neighbourhood tree

STEP 2 split into semantic components

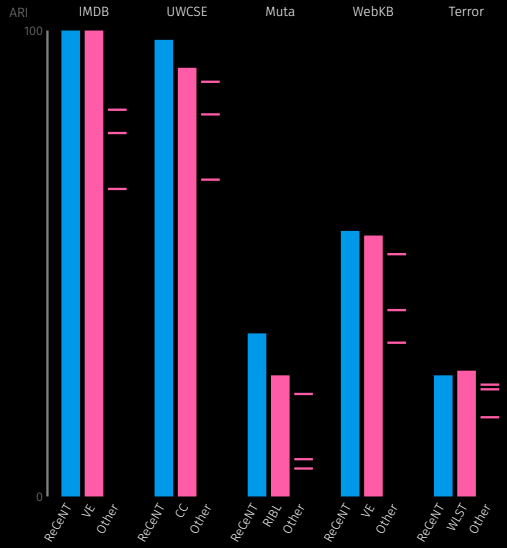


STEP 3 combine semantic components

+ + + + proximity

Evaluation

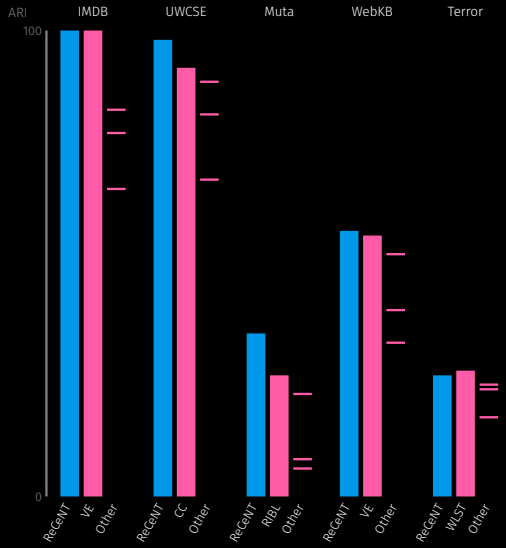
The proposed framework performs consistently well on all benchmarks



Evaluation

Main takeaway:

To reveal the structure in the relational data, one has to account for multiple sources of information/similarity



Contribution 2

Learning representations with clustering

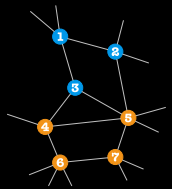
We started with a question:

How do we find groups in relational data?

Learning features = approximate symmetries

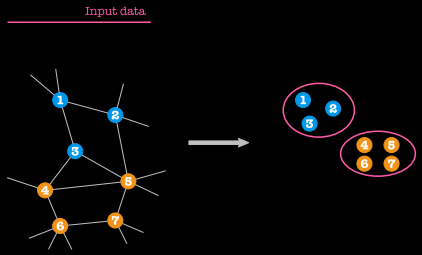
Cluster objects and relationships while altering the interpretation of similarity

Input data



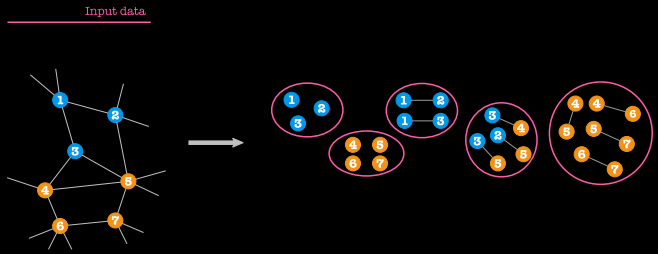
Learning features = approximate symmetries

Cluster objects and relationships while altering the interpretation of similarity



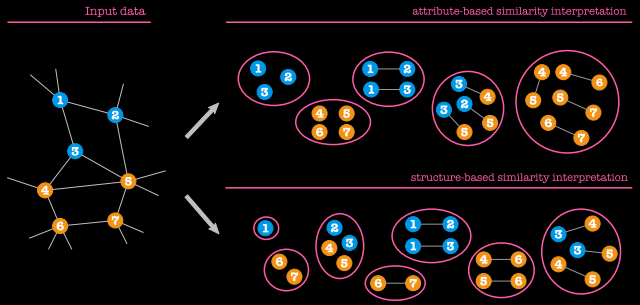
Learning features = approximate symmetries

Cluster objects and relationships while altering the interpretation of similarity



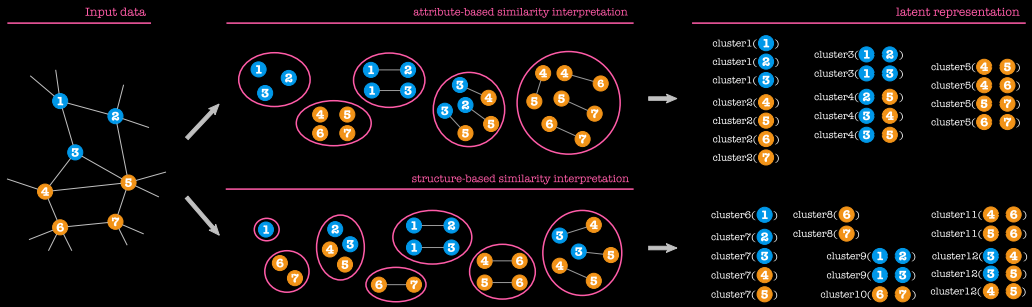
Learning features = approximate symmetries

Cluster objects and relationships while altering the interpretation of similarity



Learning features = approximate symmetries

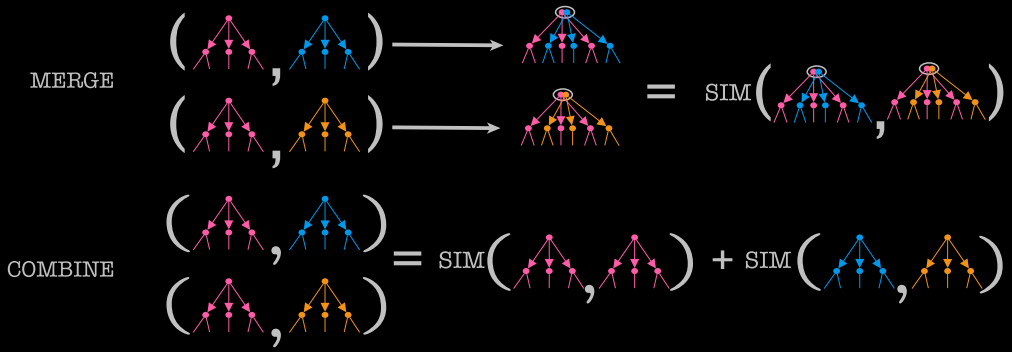
Cluster objects and relationships while altering the interpretation of similarity



A FEW MORE DETAILS

Question: *How to cluster relationships?*

- relationship = a collection of objects



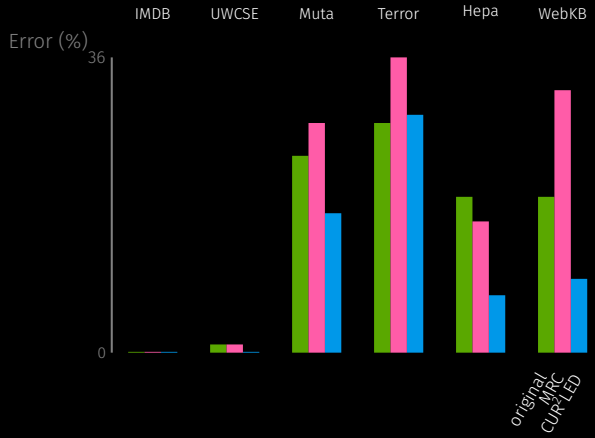
Experimental evaluation

Does CUR^{2}LED improves the performance of relational learners?

CUR^{2}LED improves the performance of relational models

and

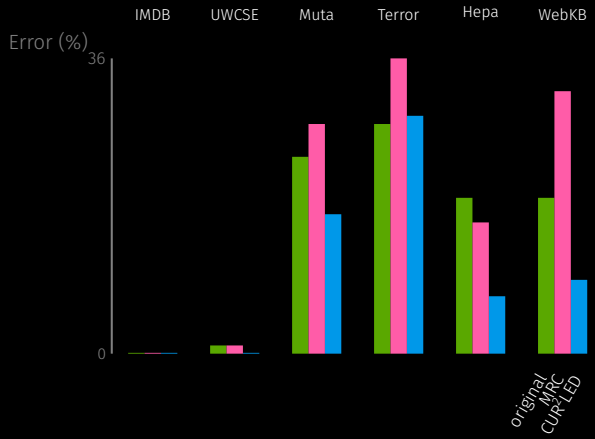
reduces the complexity



Experimental evaluation

Does CUR^{2}LED improves the performance of relational learners?

Main takeaway:
clustering is an effective way of creating latent representations, if various types of clusters are considered



Contribution 3

Auto-encoding Logic Programs

Learning representations by compression

What can we remove so that information is still meaningful?

Mathematics: multiplication compresses repetitive addition

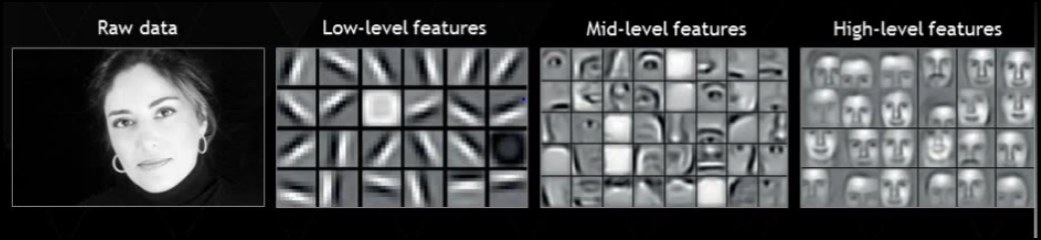
$$5 \times 2 = 2 + 2 + 2 + 2 + 2$$

Learning representations by compression

What can we remove so that information is still meaningful?

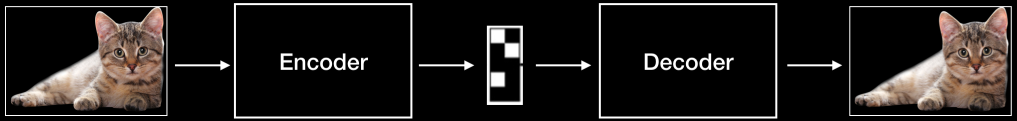
Mathematics: multiplication compresses repetitive addition

$$5 \times 2 = 2 + 2 + 2 + 2 + 2$$



Auto-encoders

- learning representations by compressing and reconstructing



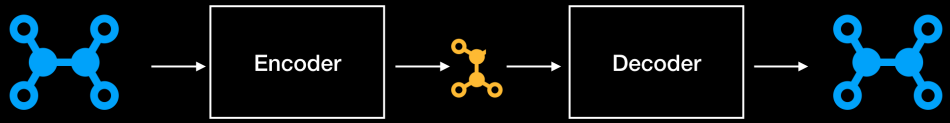
Original input

Compressed representation

Reconstructed input

Learning representations by compression

How do we compress relational data?



Original input

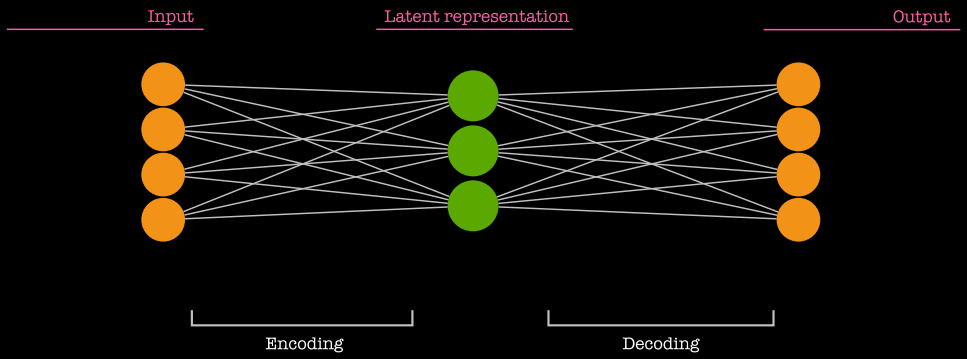
Compressed representation

Reconstructed input

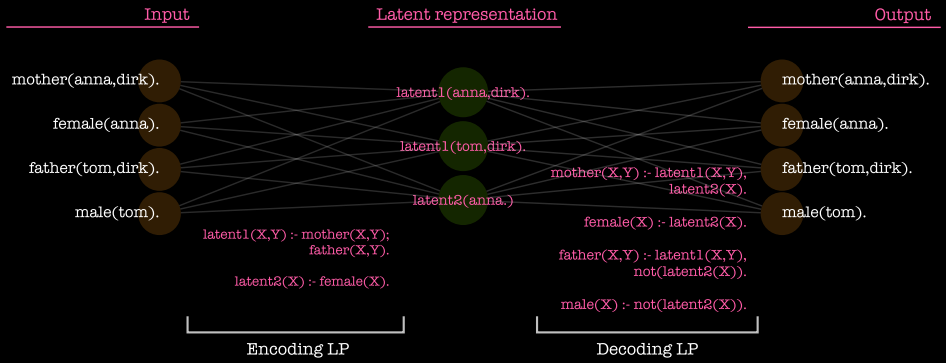
AUTO-ENCODERS

Auto-encoders: neural networks leant with back-propagation

- suitable for images and numerical data



Auto-encoding Logic Programs: use logic as a computational framework



Learning Alps in a nutshell

GIVEN a relational data

allowed logical formulas

FIND find a few formulas that reconstruct the best

built on Constraint programming technology

– effective search

– declarative

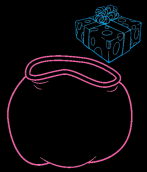
– flexible

Learning Alps in a nutshell

GIVEN a relational data

allowed logical formulas

FIND find a few formulas that reconstruct the best



Learning Alps in a nutshell

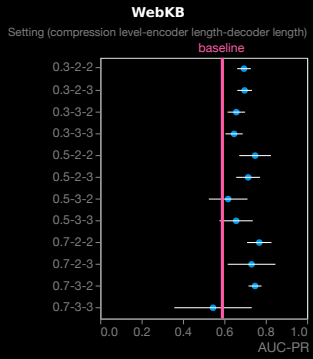
GIVEN a relational data

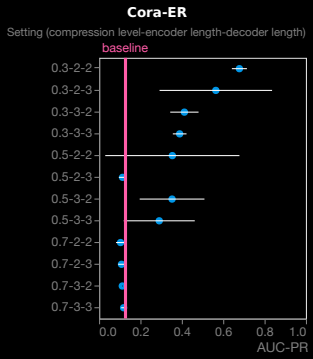
allowed logical formulas

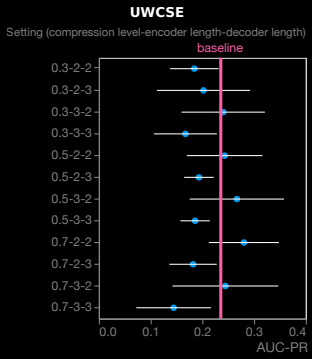
FIND find a few formulas that reconstruct the best

Experimental comparison

Do latent representation help is learning better generative models?







Contribution 4

A comparative study

Knowledge graph embeddings

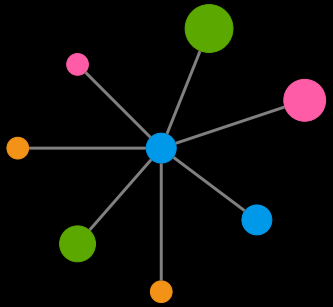
Represent objects and relations as vectors

→ [0.96, 0.45, -0.5]

→ [-0.6, 0.89, 0.05]

→ [0.12, 0.47, -0.47]

→ [0.96, 0.07, -0.12]



Knowledge graph embeddings

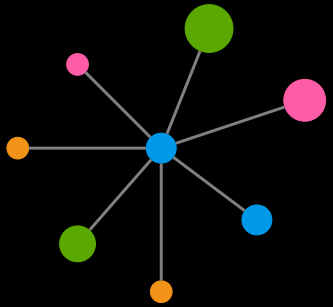
Represent objects and relations as vectors

existing edges get a high score

score(, ,)

non-existing edges get a low score

score(, ,)



Knowledge graph embeddings

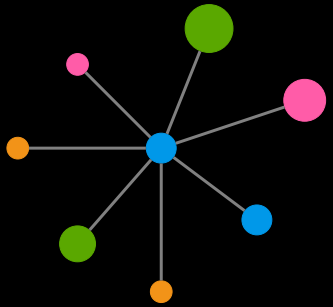
Represent objects and relations as vectors

existing edges get a high score

score(●, ▬, ●)

non-existing edges get a low score

score(●, ▬, ●)



score can be any operation between vectors

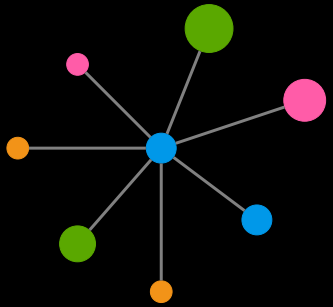
score(●, ▬, ●) = − || ● + ▬ − ● ||

Knowledge graph embeddings

Represent objects and relations as vectors

Back to Excel tables!

[0.96, 0.45, -0.5]
[-0.6, 0.89, 0.05]
[0.12, 0.47, -0.47]



Sub-symbolic methods

- efficient (fast)
- large data
- 'simple' framework

Symbolic methods

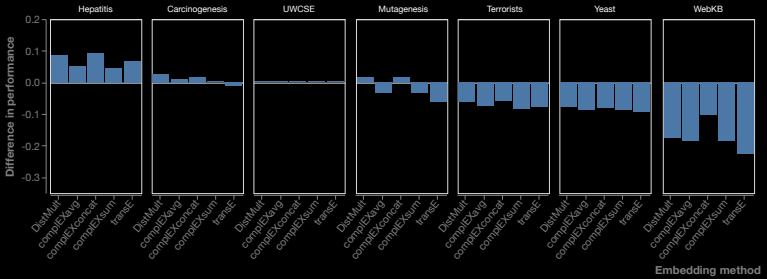
- expressive
- flexible
- interpretable

What do we gain/loose with embeddings?

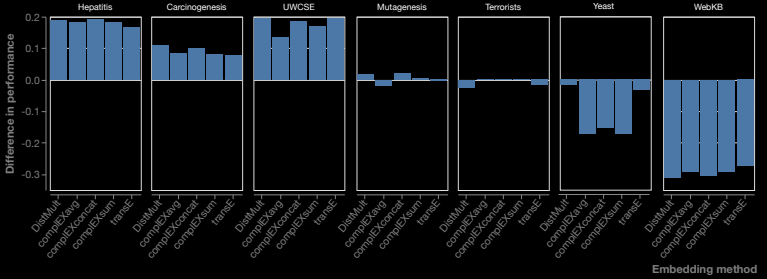
Performance comparison on standard benchmarks

Embeddings perform better when data is curated

Decision trees

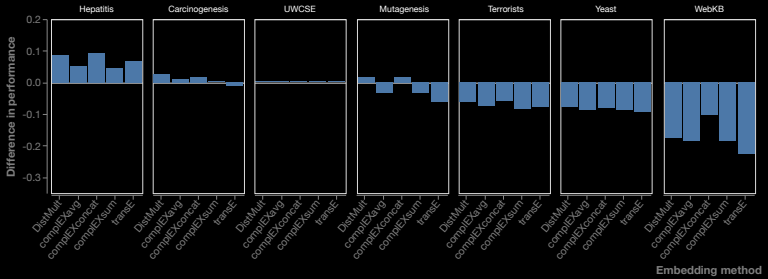


SVMs



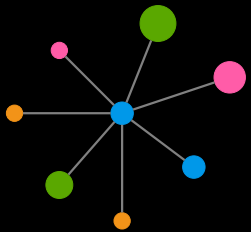
Embeddings perform better when data is curated

Decision trees

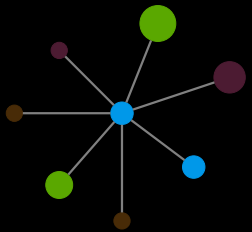


The proportion of *predictive information* indicative of the performance

Input data

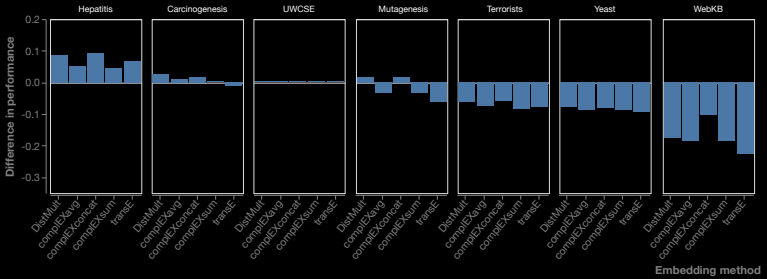


Predictive data
= used for
prediction by a
symbolic model



Embeddings perform better when data is curated

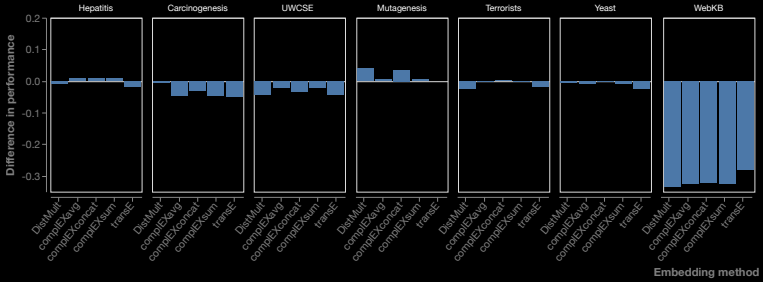
Decision trees



Dataset	Model complexity	Dataset properties		
		Attributes	Relations	Neighbourhood
Hepatitis	15	100 %	66 %	70 %
Carcinogenesis	24	–	100 %	72 %
UWCSE	35.5	75 %	80 %	0 %
Terrorists	17.65	35 %	100 %	0 %
Mutagenesis	9.4	100 %	0 %	0 %
WebKB	38	4 %	50 %	1 %
Yeast	188.4	18 %	100 %	–

Embeddings are good at estimating similarity

Similarity-based learning



Main takeaway:

Embeddings are suitable for curated data and similarity estimation for relational objects

Conclusion

Conclusions

PROBLEM How do we learn symbolic latent representation for relational data?

CONTRIBUTION 1 A versatile relational clustering framework
CONTRIBUTION 2 Learning representations through symmetries
CONTRIBUTION 3 Auto-encoding Logic Programs
CONTRIBUTION 4 A comparison of symbolic and sub-symbolic approaches