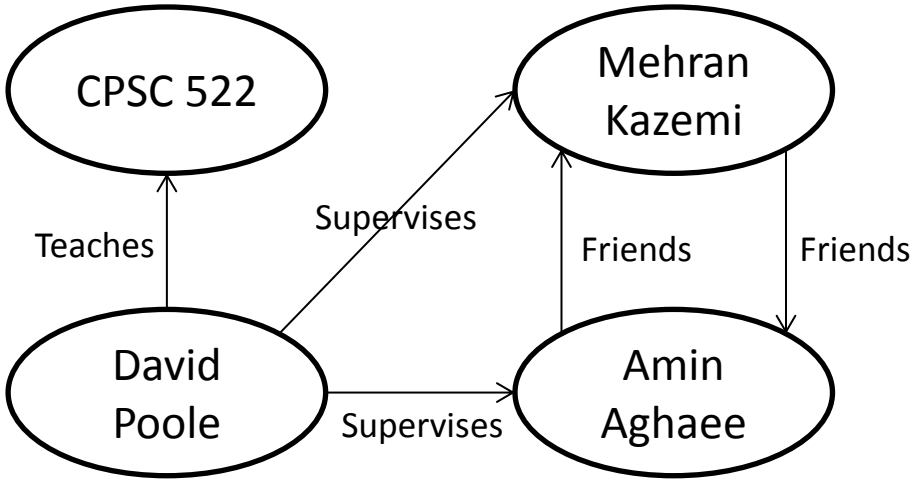
Tensor Factorization
for
Knowledge Graph Completion

Mehran Kazemi

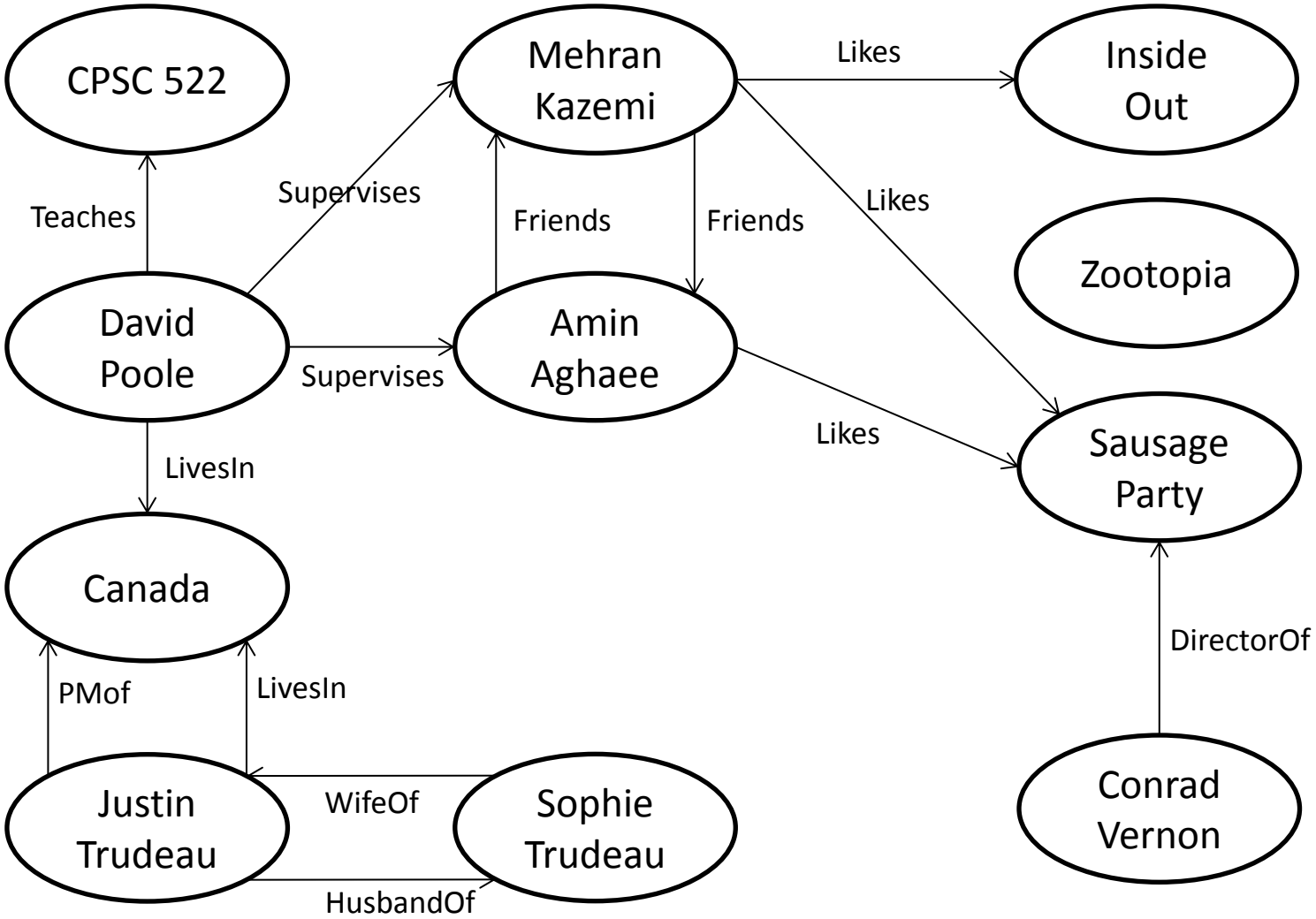
Guest Lecture for CPSC-522 at CS@UBC

Winter 2018

Knowledge Graph

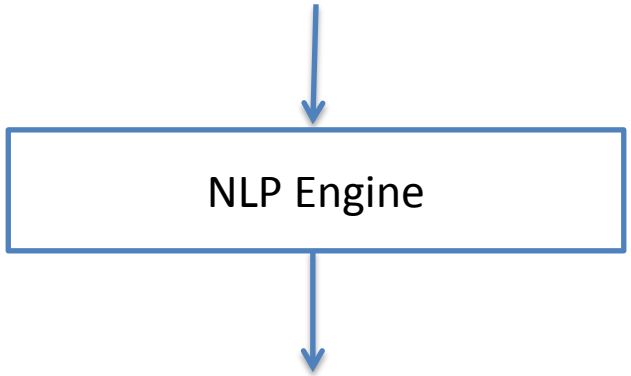


Knowledge Graph



Knowledge Graph Construction

Justin Trudeau is a Canadian politician serving as the 23rd and current prime minister of Canada since 2015 and leader of the liberal party since 2013.



- (Justin Trudeau, Nationality, Canada)
- (Justin Trudeau, PMof, Canada)
- …

Knowledge Graph Construction

Justin Trudeau is a Canadian politician serving as the 23rd and current prime minister of Canada since 2015 and leader of the liberal party since 2013.

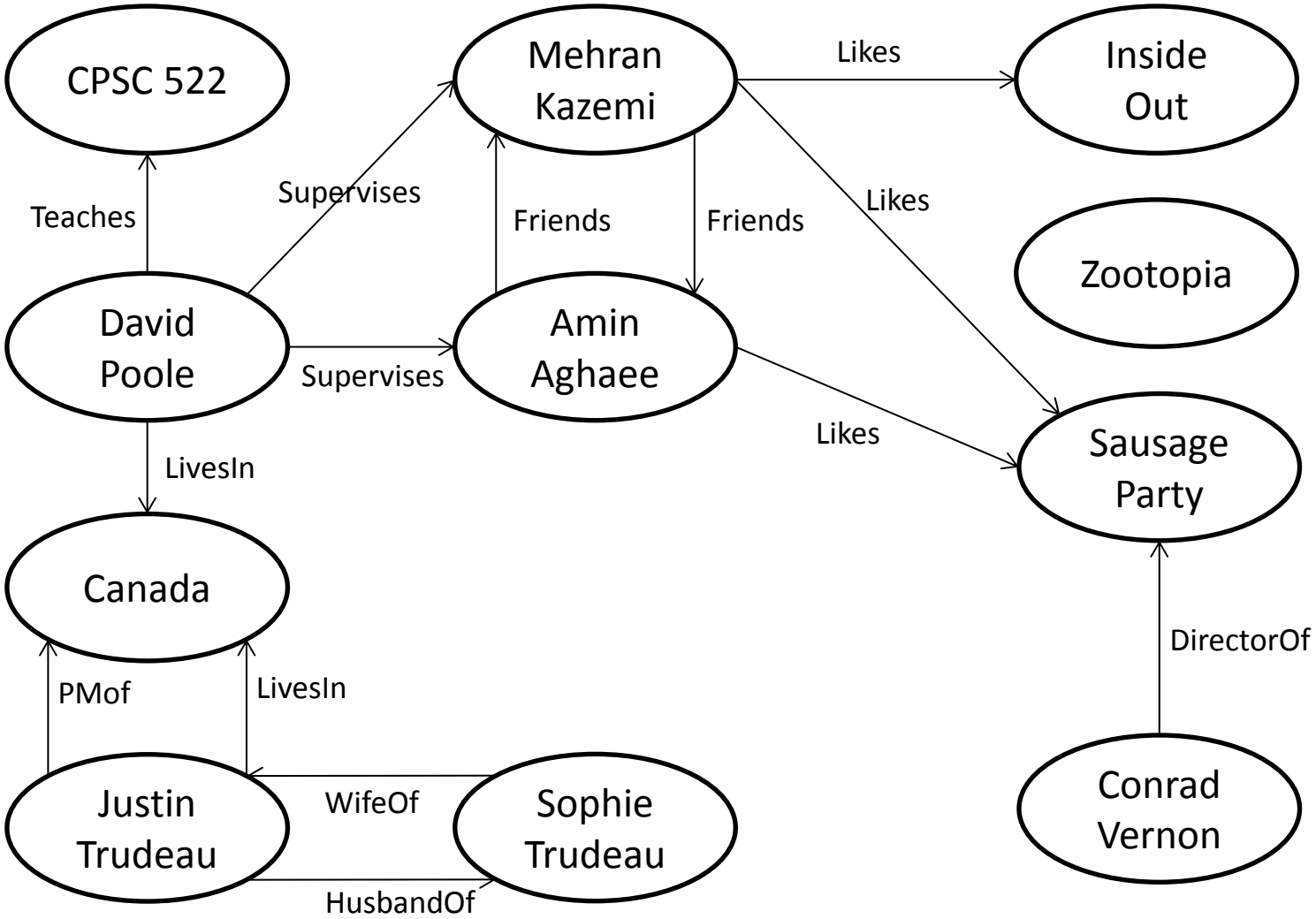
NLP Engine

- (Justin Trudeau, Nationality, Canada)
- (Justin Trudeau, PMof, Canada)
- …

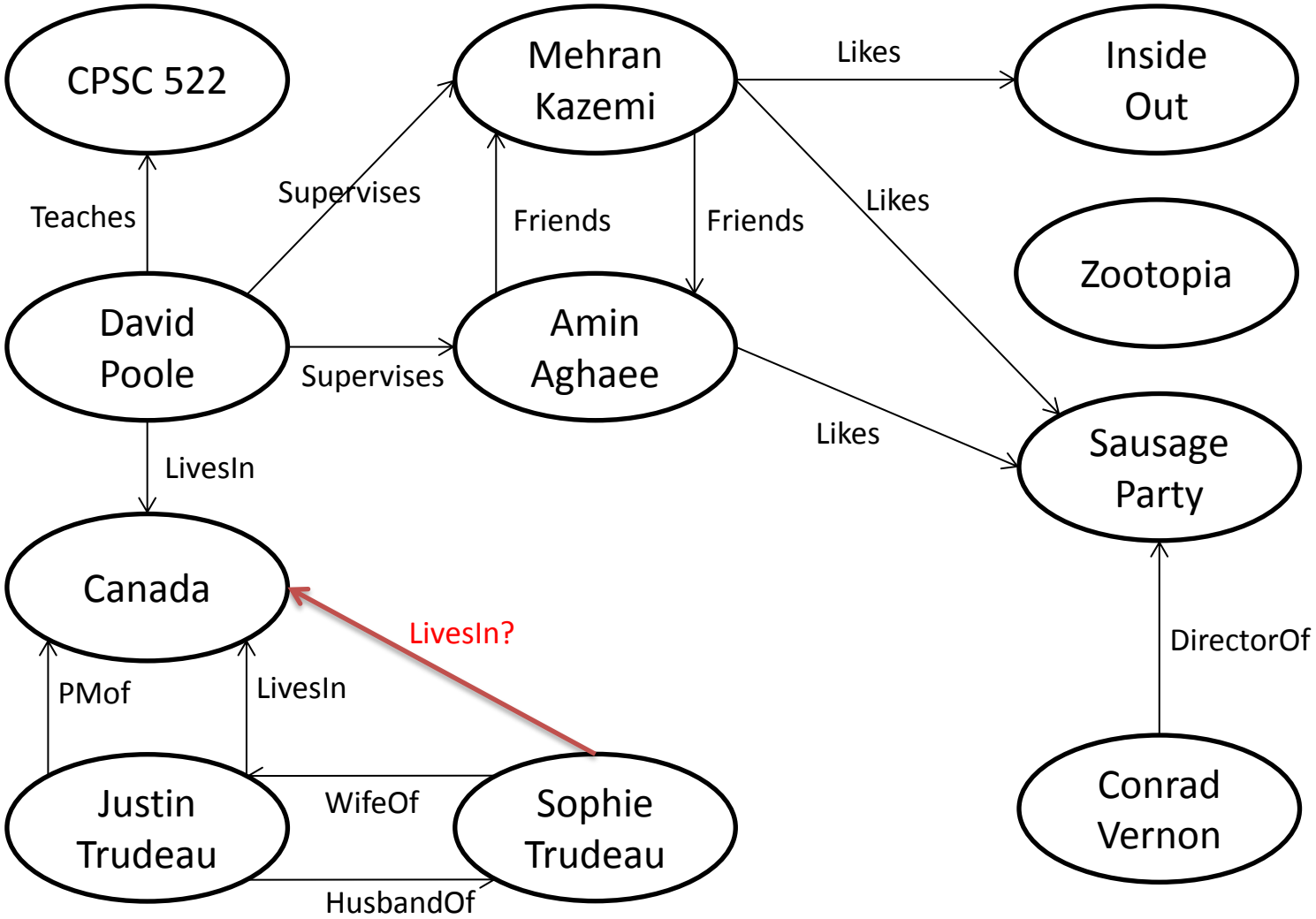
Computer Vision Engine

- (Justin Trudeau, HusbandOf, Sophie Trudeau)
- (Sophie Trudeau, WifeOf, Justin Trudeau)

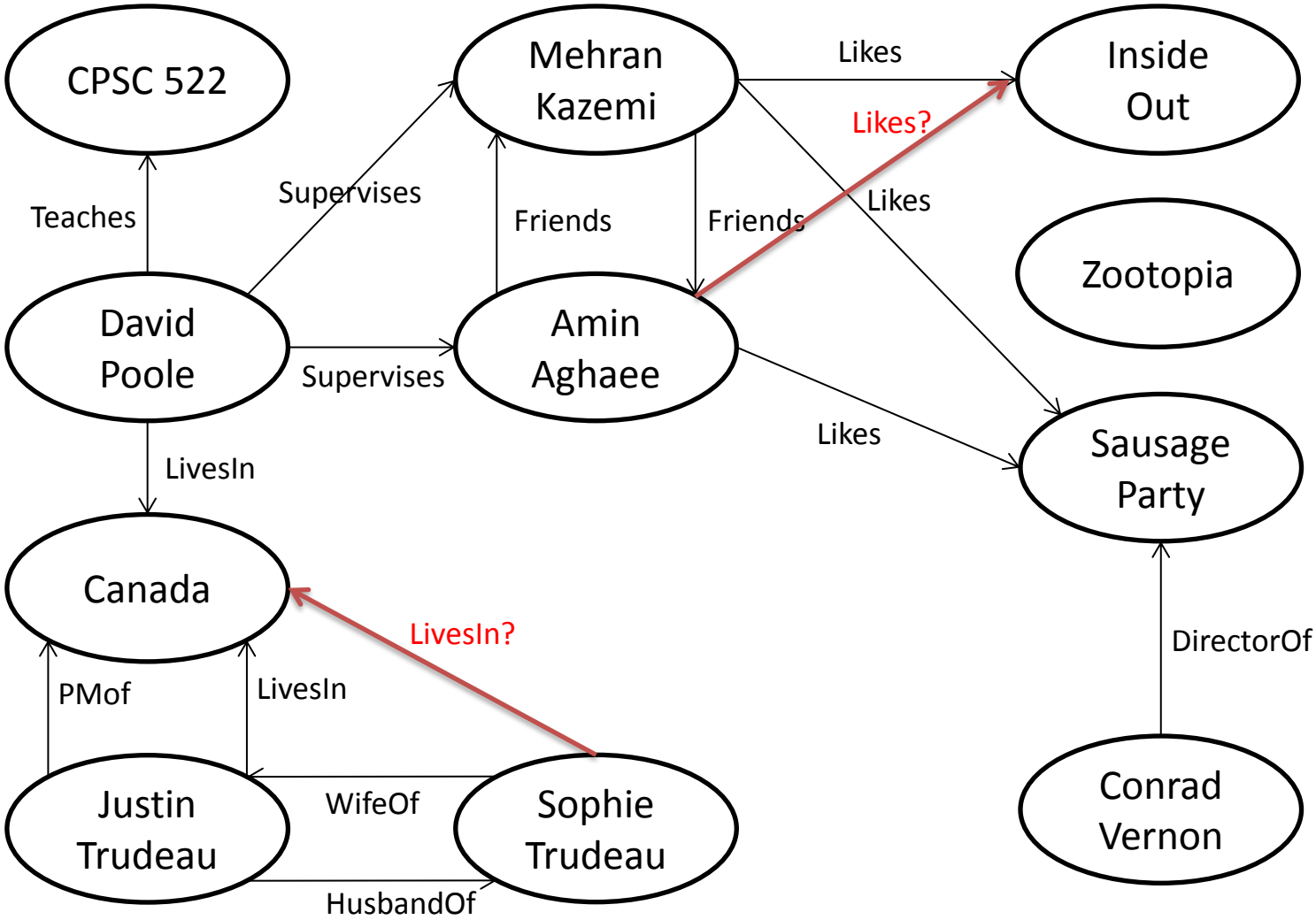
Knowledge Graphs are Incomplete



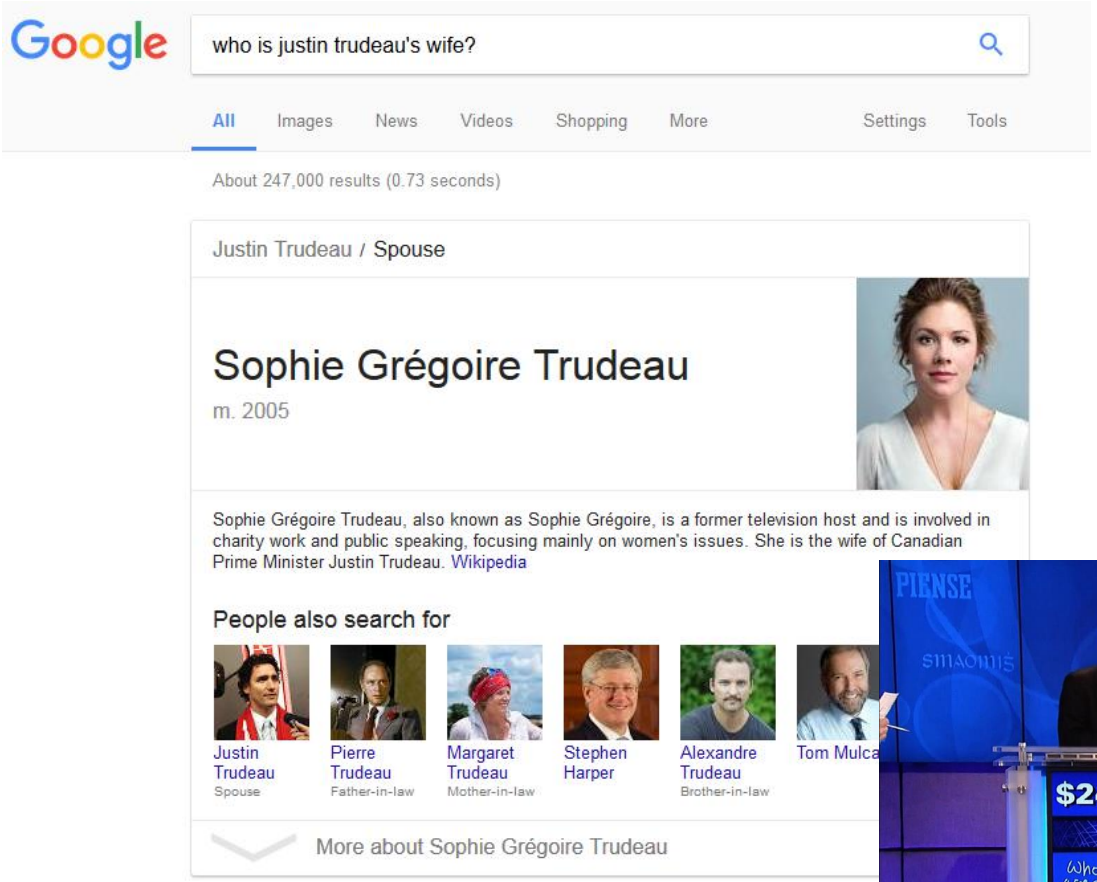
Knowledge Graph Completion



Knowledge Graph Completion

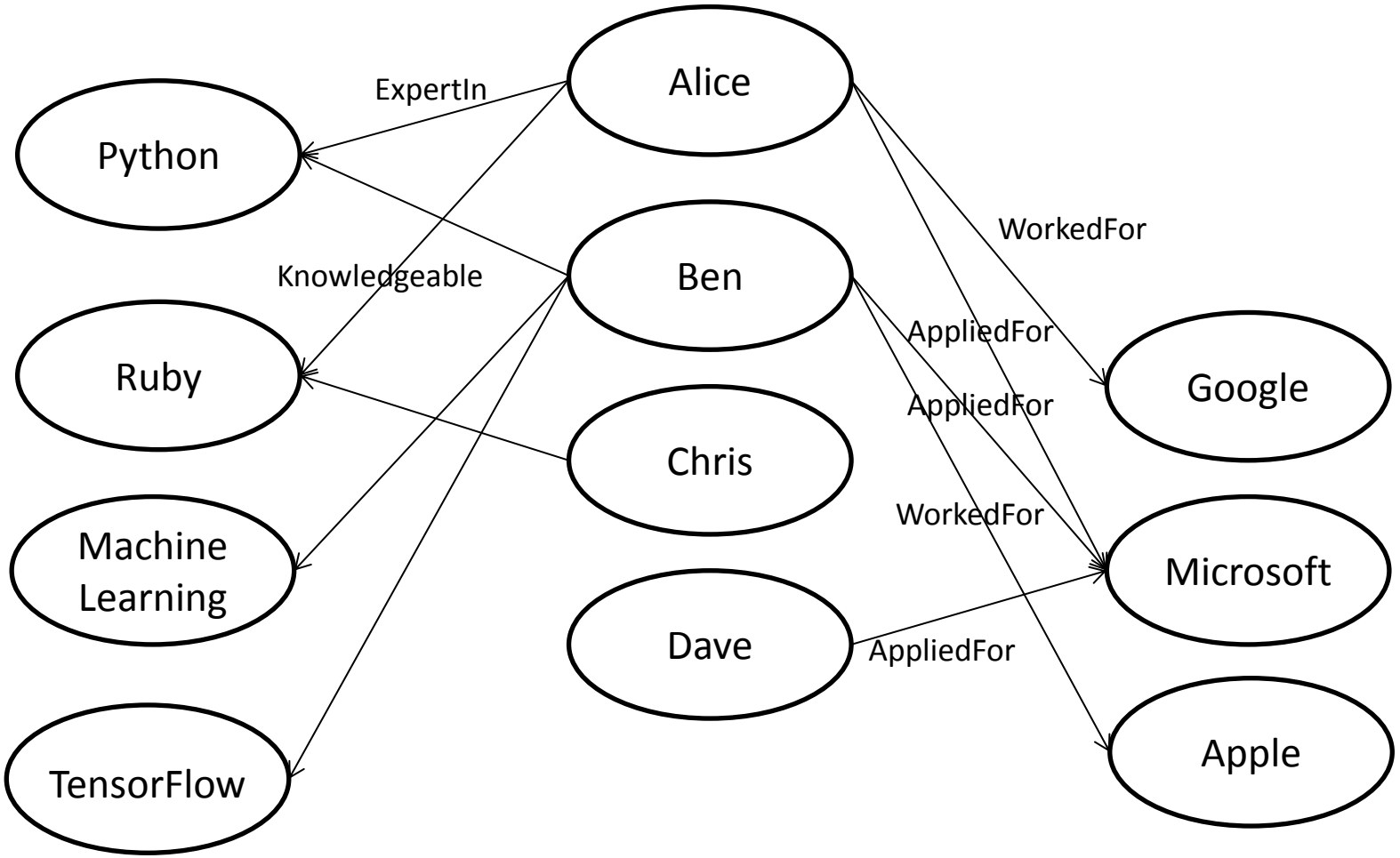


Applications: QA



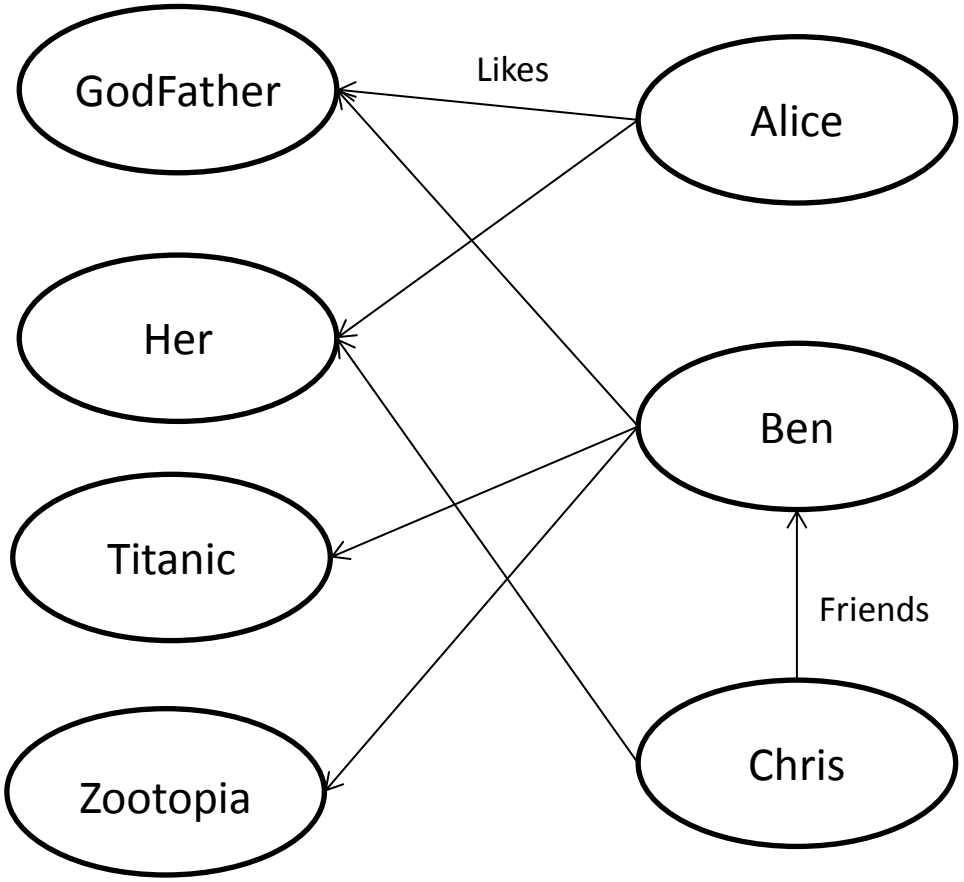
Applications: Recruiting

Which applicant should Microsoft hire?



Applications: Recom. Systems

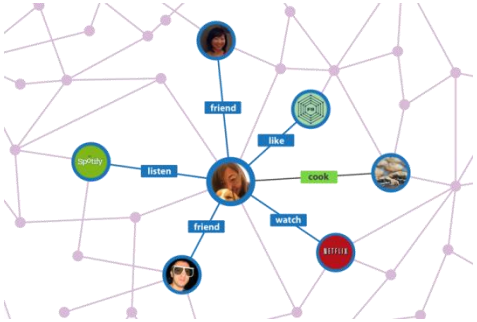
Which movie should we recommend to Chris?

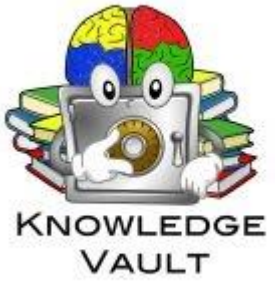


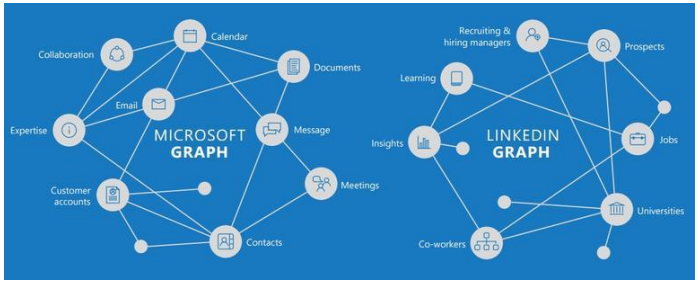
Applications: Predicting the Elections

Knowledge Graphs & Industry

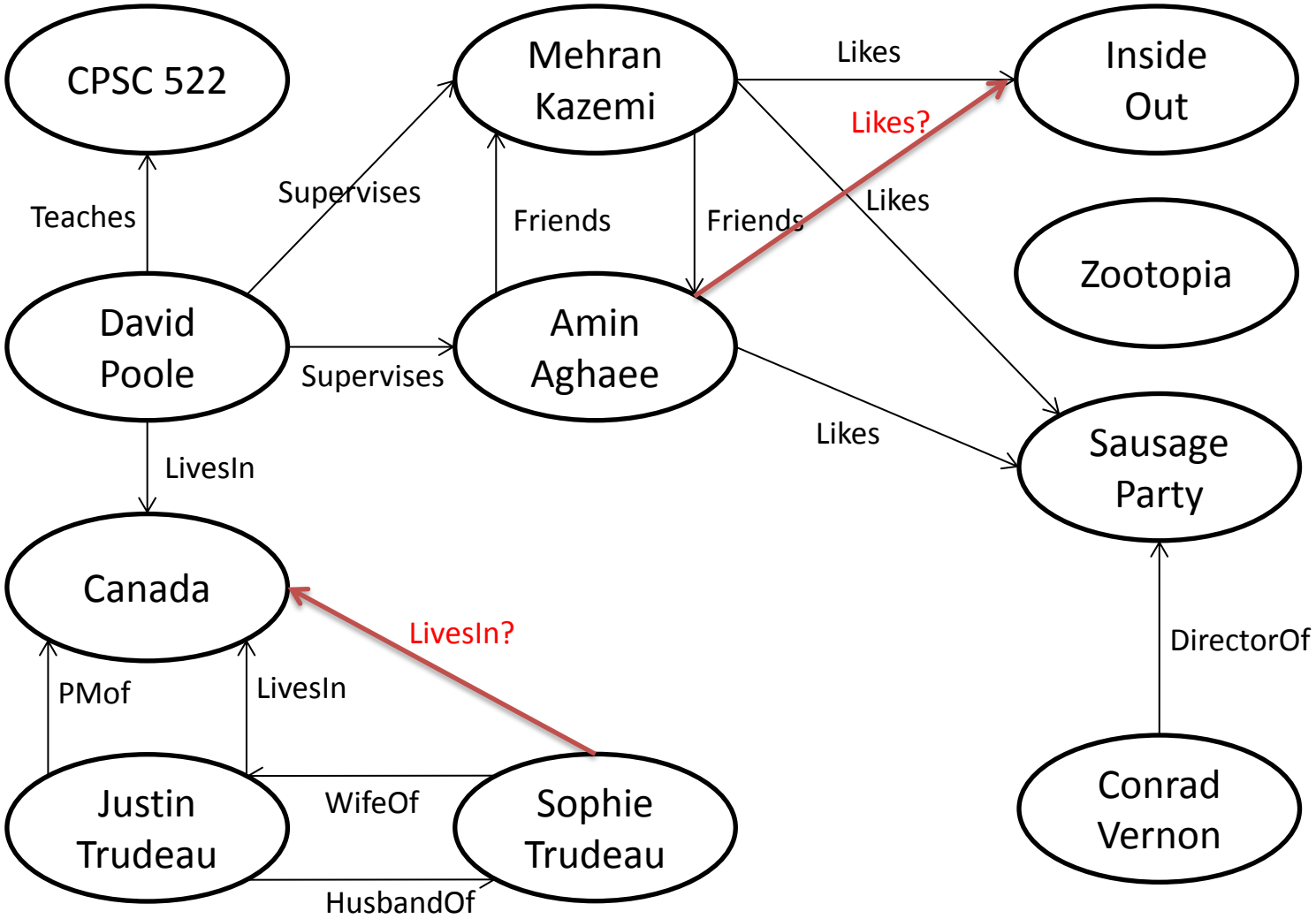
- Google Knowledge Vault
- Amazon Product Graph
- Facebook Graph API
- IBM Watson
- Microsoft Satori
- LinkedIn Knowledge Graph
- Yandex Object Answer
- …







How to Predict New Links?



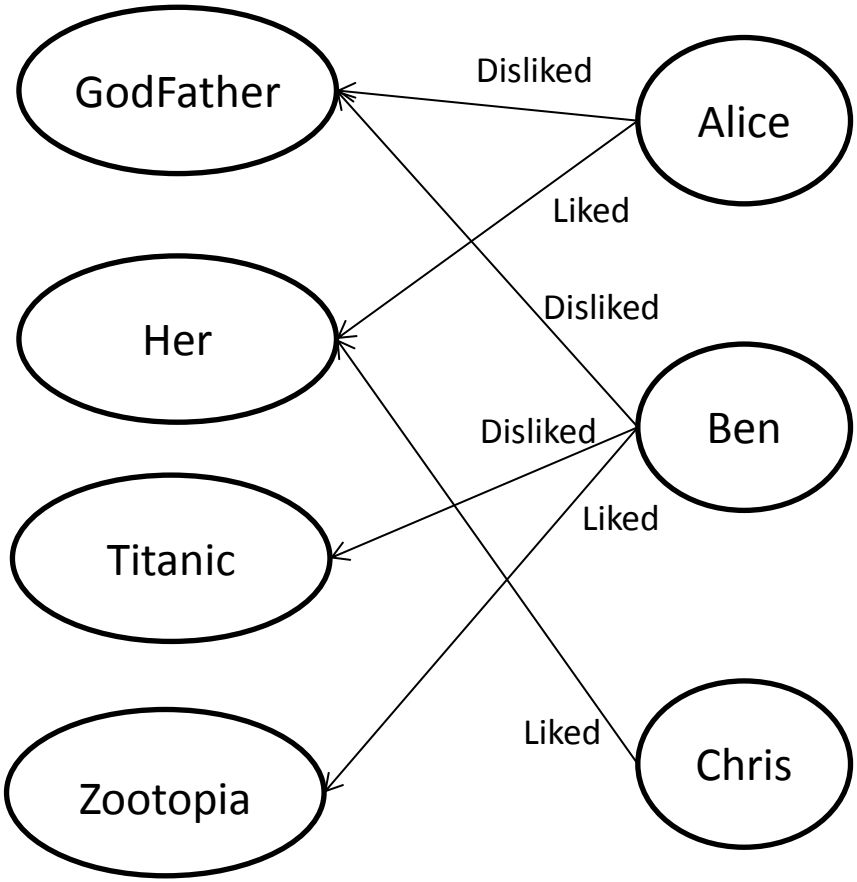
How to Predict New Links?

- Weighted rules:
  - 0.9: $LivesIn(X, L) \wedge HusbandOf(X, Z) \Rightarrow LivesIn(Z, L)$
    - Markov Logic Networks
    - ProbLog
    - Probabilistic Soft Logic
    - Relational Logistic Regression

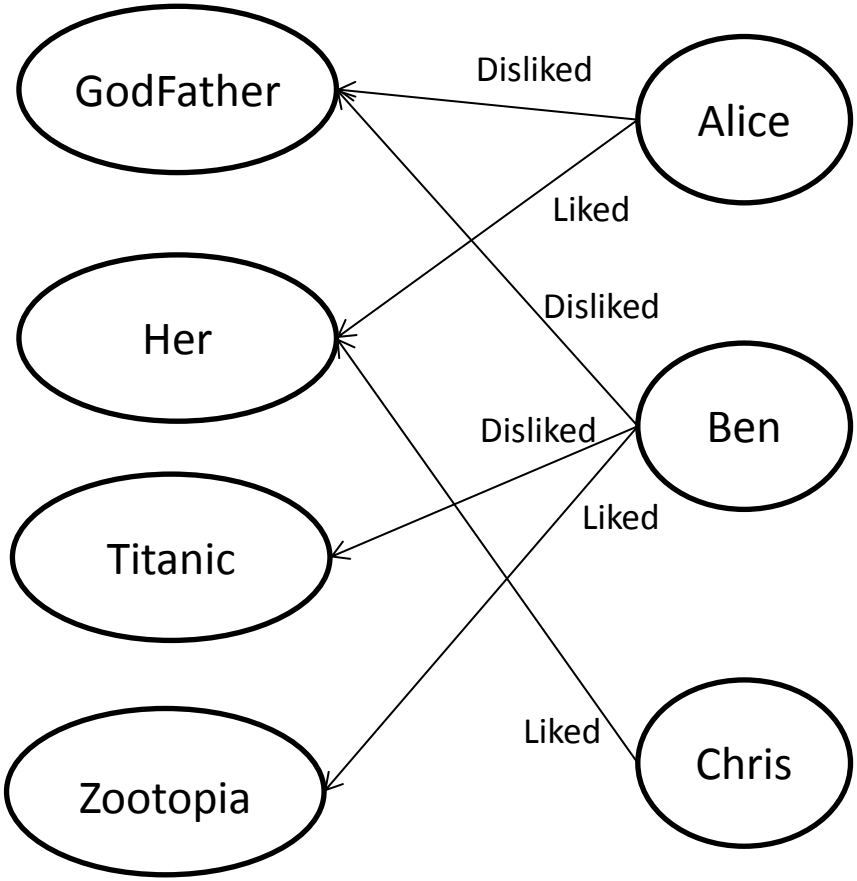
How to Predict New Links?

- Weighted rules:
  - 0.9: $LivesIn(X, L) \wedge HusbandOf(X, Z) \Rightarrow LivesIn(Z, L)$
    - Markov Logic Networks
    - ProbLog
    - Probabilistic Soft Logic
    - Relational Logistic Regression
- Tensor Factorization:
  - Learn embeddings for objects and relationships
  - Learn a probabilistic function from embeddings to whether a link exists between two entities or not

Tensor Factorization



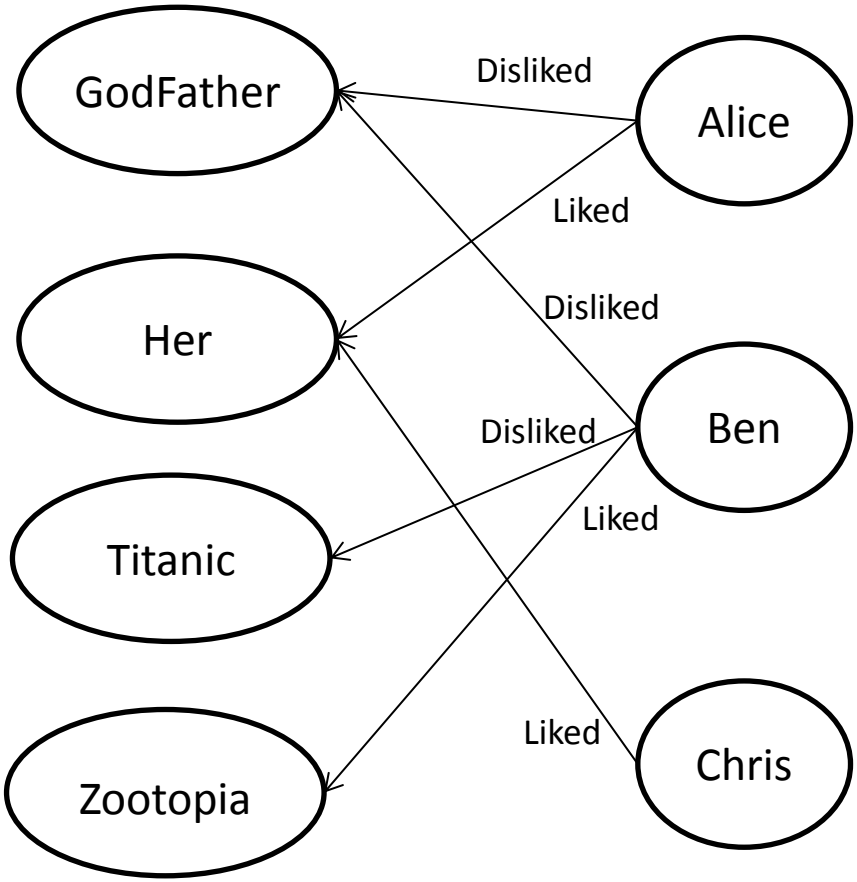
Tensor Factorization



	Likes long movies	Likes action	Likes romance	Likes X	Likes Y
Alice:	0.1	0	-3	1.1	-1

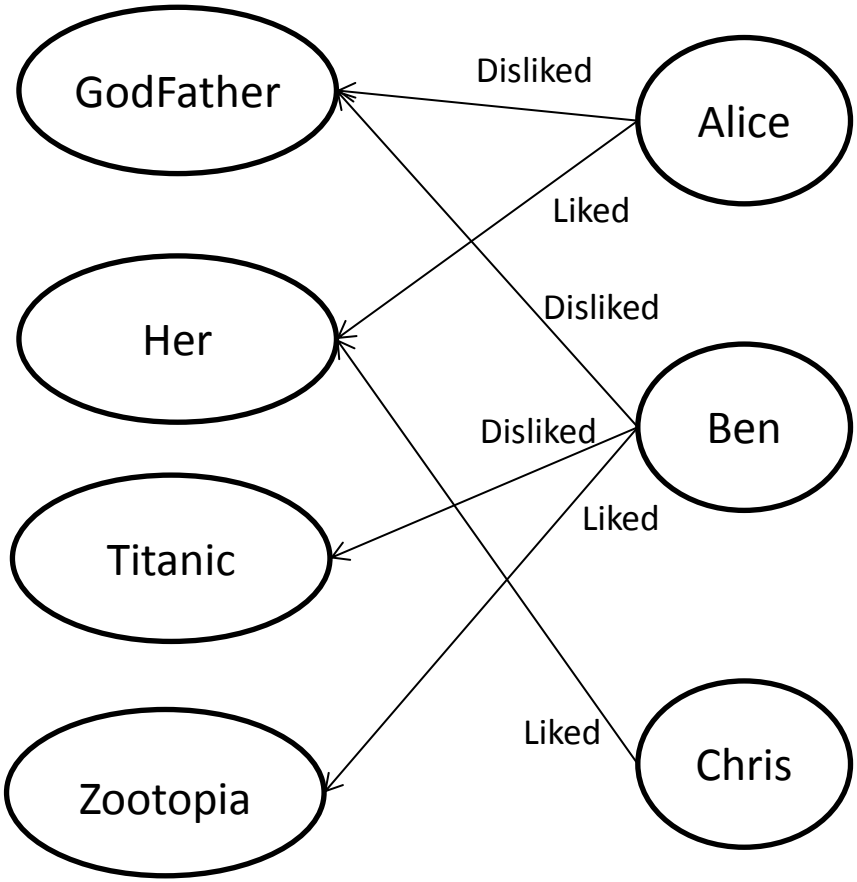
	Is long	Has action	Has romance	Has X	Has Y
Titanic:	0.5	-2	5	1	-2

Tensor Factorization



	Likes long movies	Likes action	Likes romance	Likes X	Likes Y
Alice:	0.1	0	-3	1.1	-1
	0.05	0	-15	1.1	2
Titanic:	0.5	-2	5	1	-2
	Is long	Has action	Has romance	Has X	Has Y

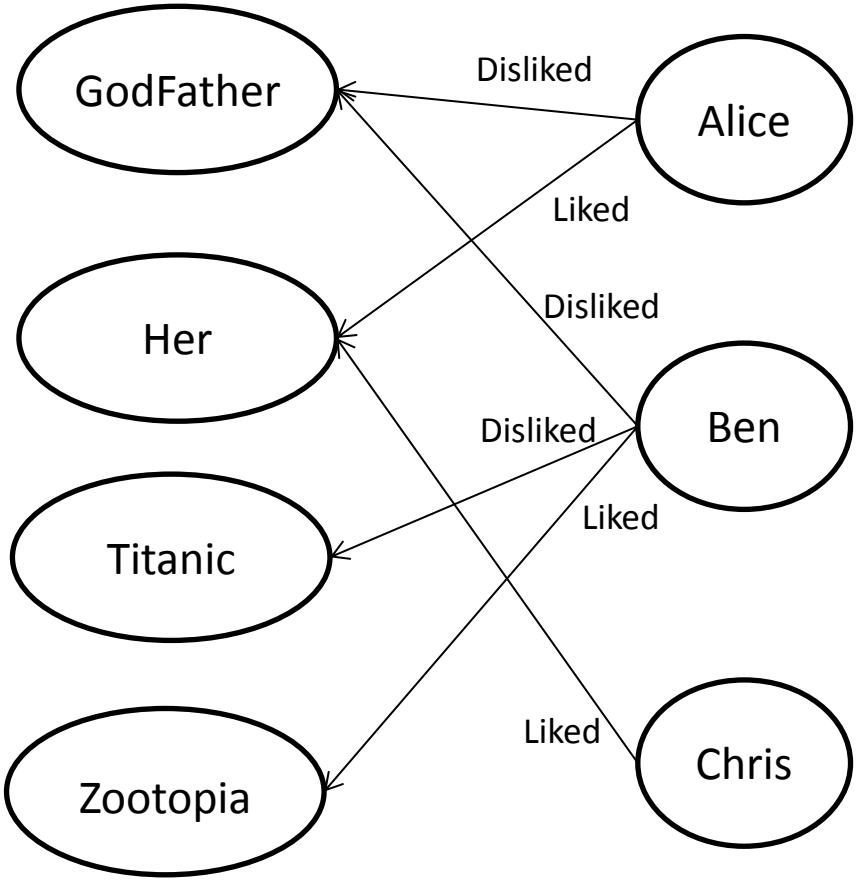
Tensor Factorization

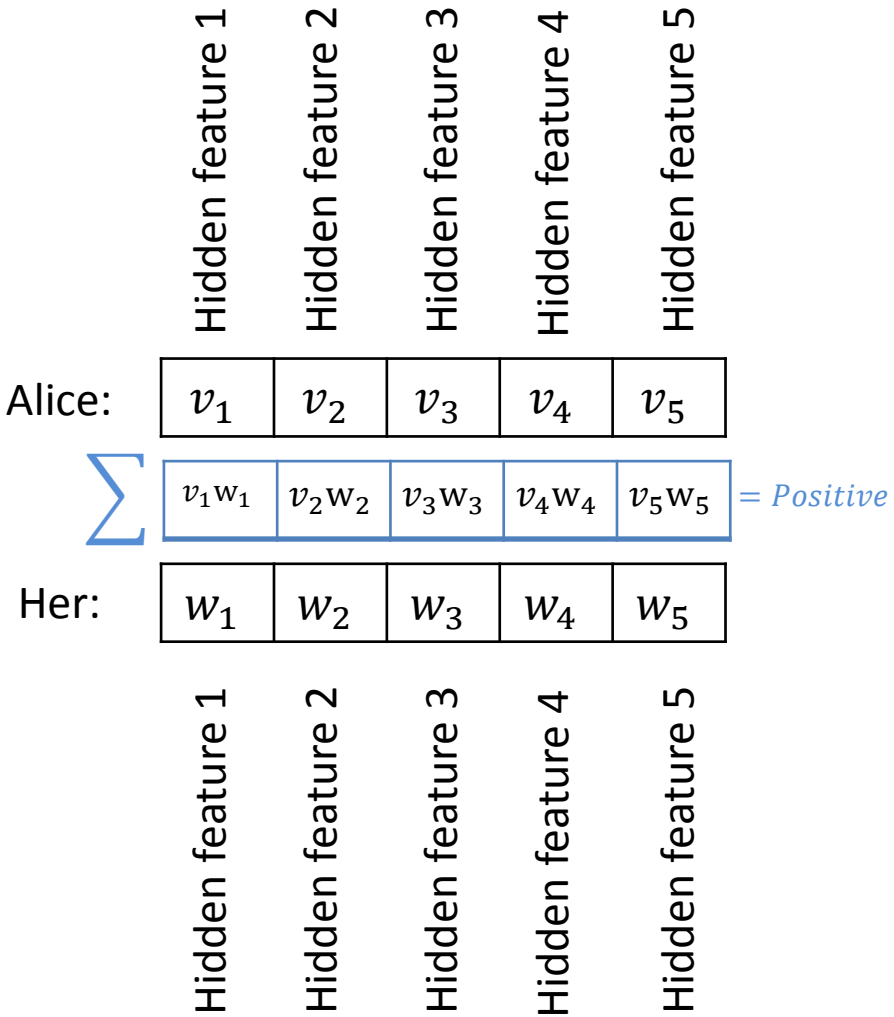


	Likes long movies	Likes action	Likes romance	Likes X	Likes Y
Alice:	0.1	0	-3	1.1	-1
$\sum$	0.05	0	-15	1.1	2
Titanic:	0.5	-2	5	1	-2
	Is long	Has action	Has romance	Has X	Has Y

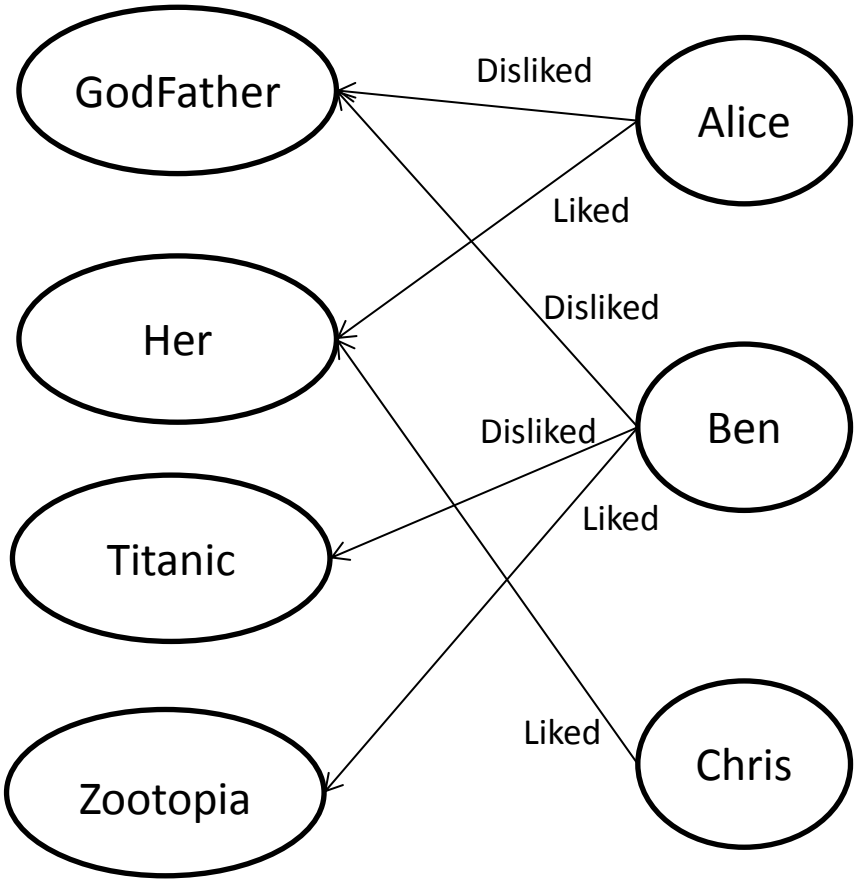
$= -11.85$

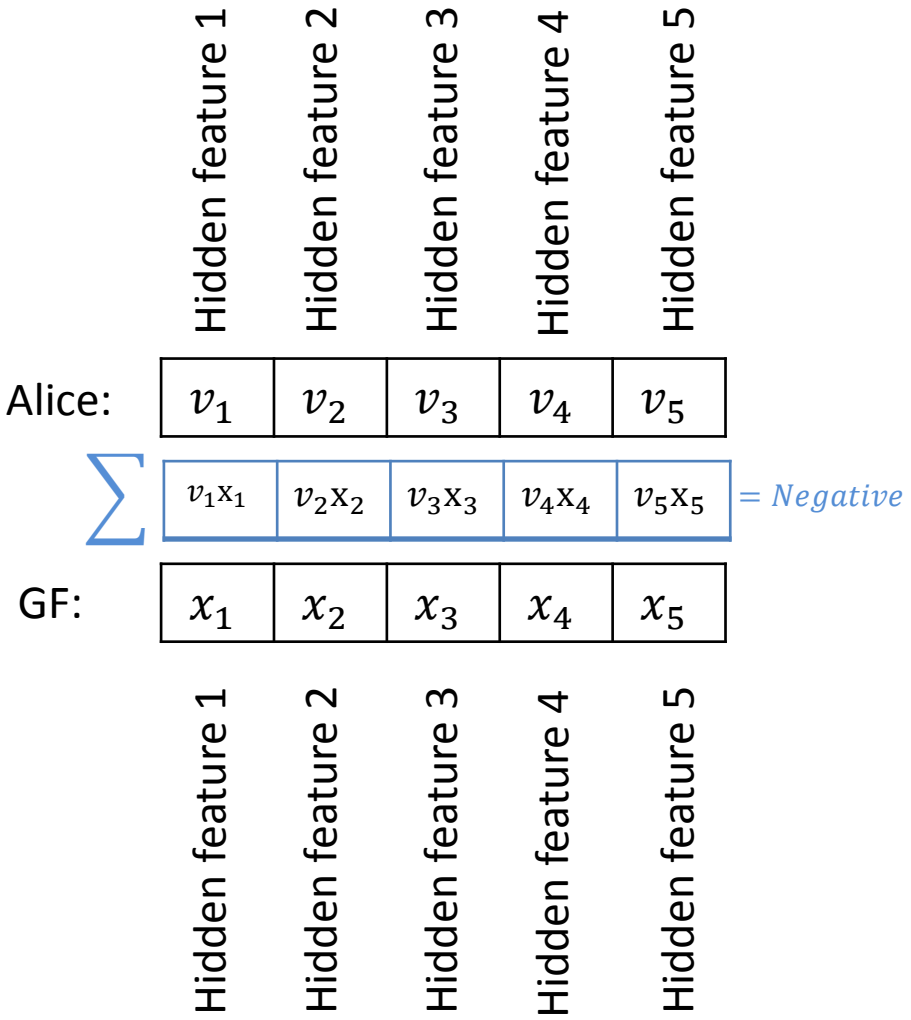
What if we don't have the features?



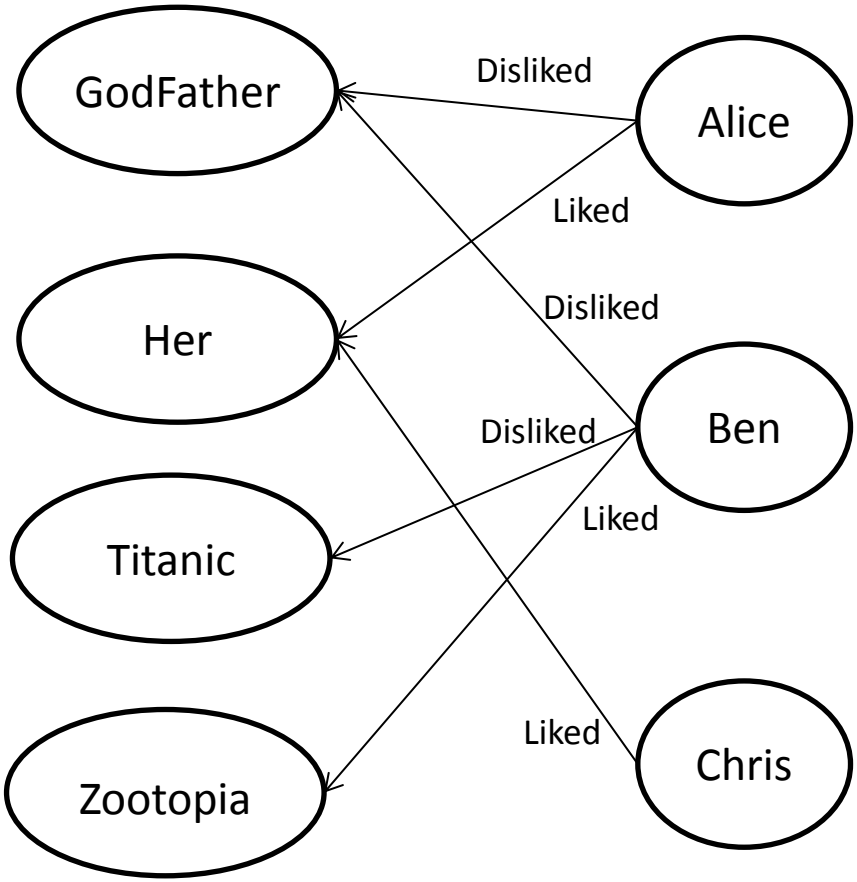


What if we don't have the features?





Learning Hidden Features

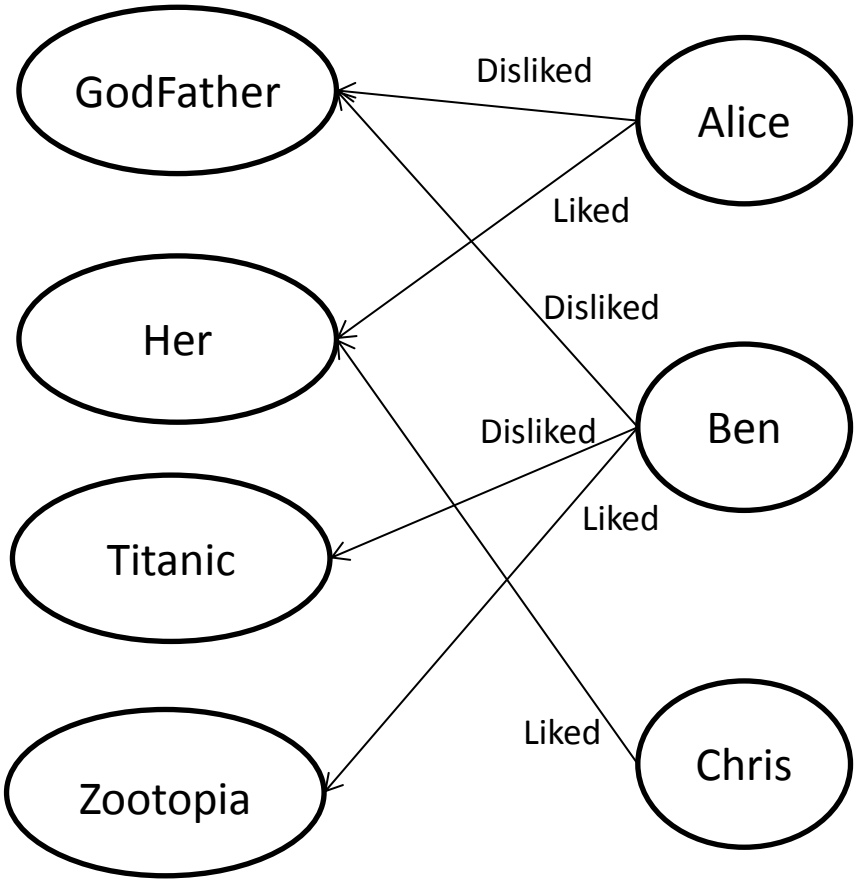


Random Initialization

Alice:	1.1	−2	1	1	−1
Ben:	−1	−2	3	1	1.2
Chris:	3.1	−1	−0.2	0.2	0.3
GF:	1.5	0.2	−2	−9	1
Her:	0	0.1	−1	1	2

…

Learning Hidden Features



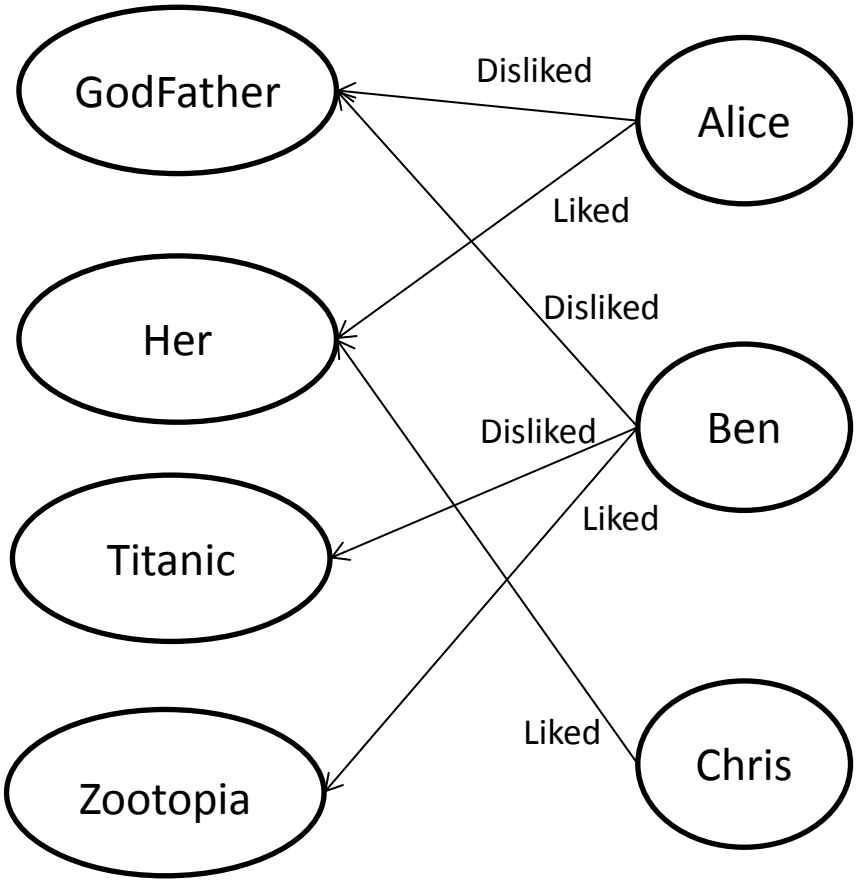
Random Initialization

Alice:	1.1	−2	1	1	−1
Ben:	−1	−2	3	1	1.2
Chris:	3.1	−1	−0.2	0.2	0.3
GF:	1.5	0.2	−2	−9	1
Her:	0	0.1	−1	1	2

...

Alice:	1.1	−2	1	1	−1	
$\sum$	0	−0.2	−1	1	−2	$= -2.2$
Her:	0	0.1	−1	1	2	

Learning Hidden Features



Random Initialization

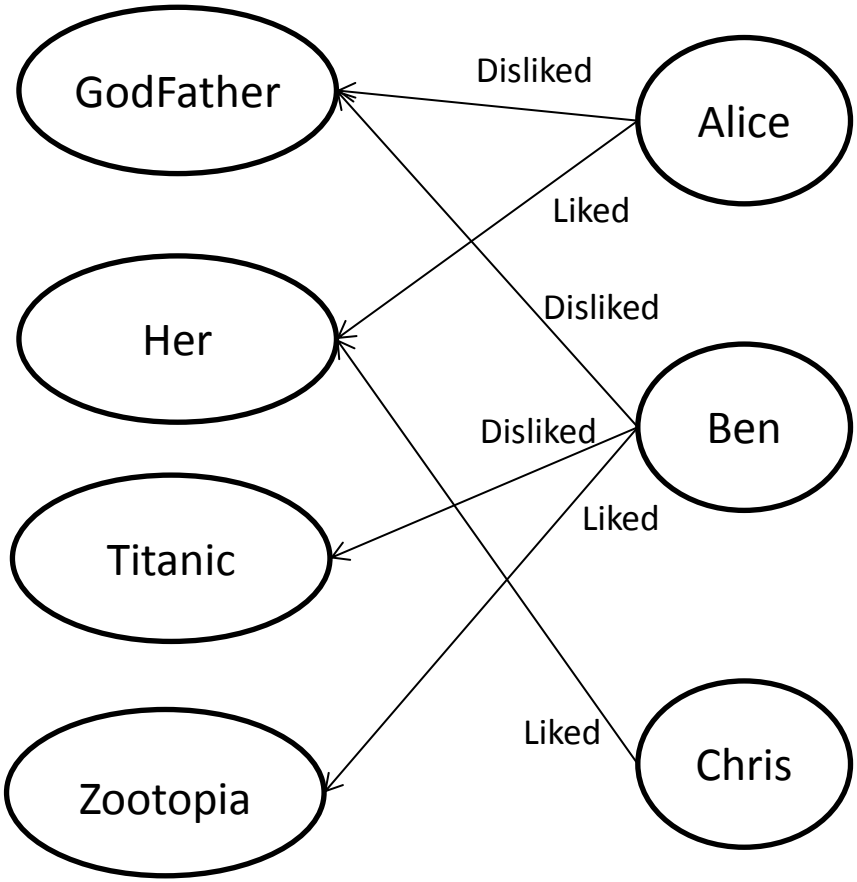
Alice:	1.1	−2	1	1	−1
Ben:	−1	−2	3	1	1.2
Chris:	3.1	−1	−0.2	0.2	0.3
GF:	1.5	0.2	−2	−9	1
Her:	0	0.1	−1	1	2

…

Alice:	1.2	−1.9	0.9	1.1	−0.9

Her:	0.1	0	−0.9	1.1	1.9

Learning Hidden Features



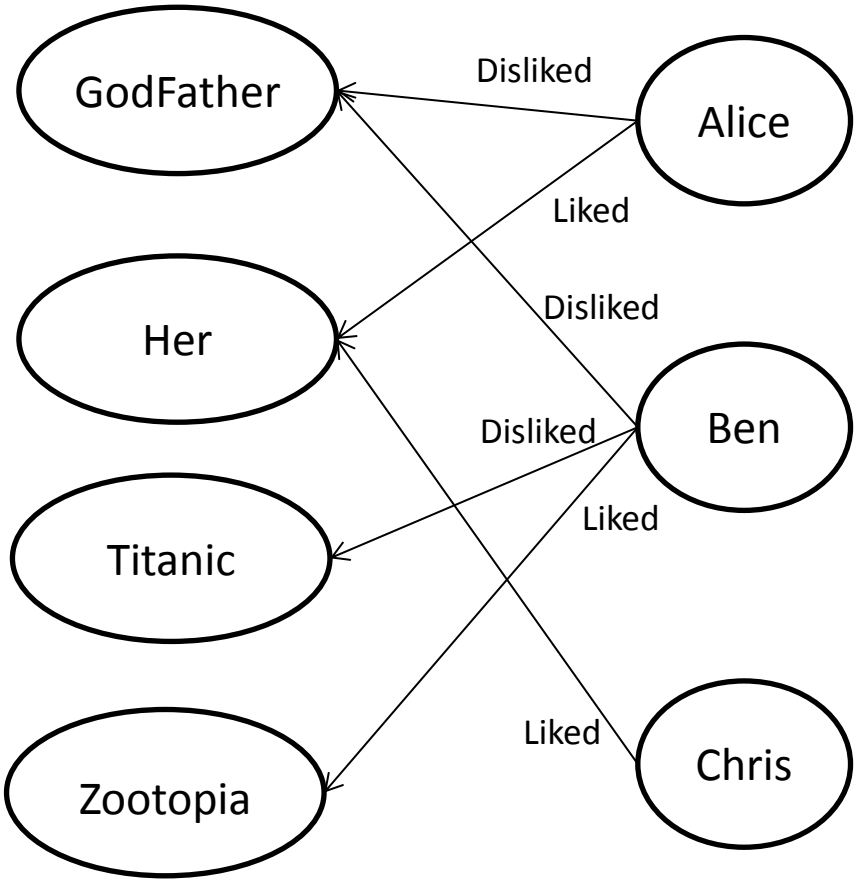
Random Initialization

Alice:	1.1	−2	1	1	−1
Ben:	−1	−2	3	1	1.2
Chris:	3.1	−1	−0.2	0.2	0.3
GF:	1.5	0.2	−2	−9	1
Her:	0	0.1	−1	1	2

...

Alice:	1.2	−1.9	0.9	1.1	−0.9	
$\sum$	0.12	0	−0.81	1.21	−1.71	$= -1.19$
Her:	0.1	0	−0.9	1.1	1.9	

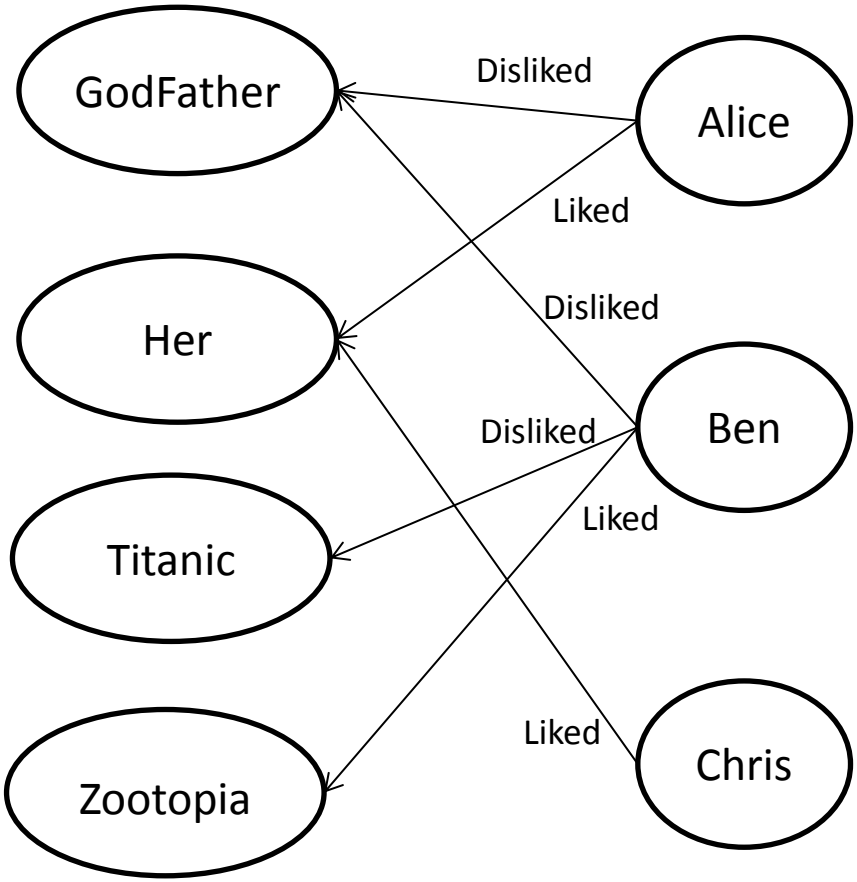
Learning Hidden Features



Alice:	1.2	−1.9	0.9	1.1	−0.9
Ben:	−1	−2	3	1	1.2
Chris:	3.1	−1	−0.2	0.2	0.3
GF:	1.5	0.2	−2	−9	1
Her:	0.1	0	−0.9	1.1	1.9

…

Learning Hidden Features

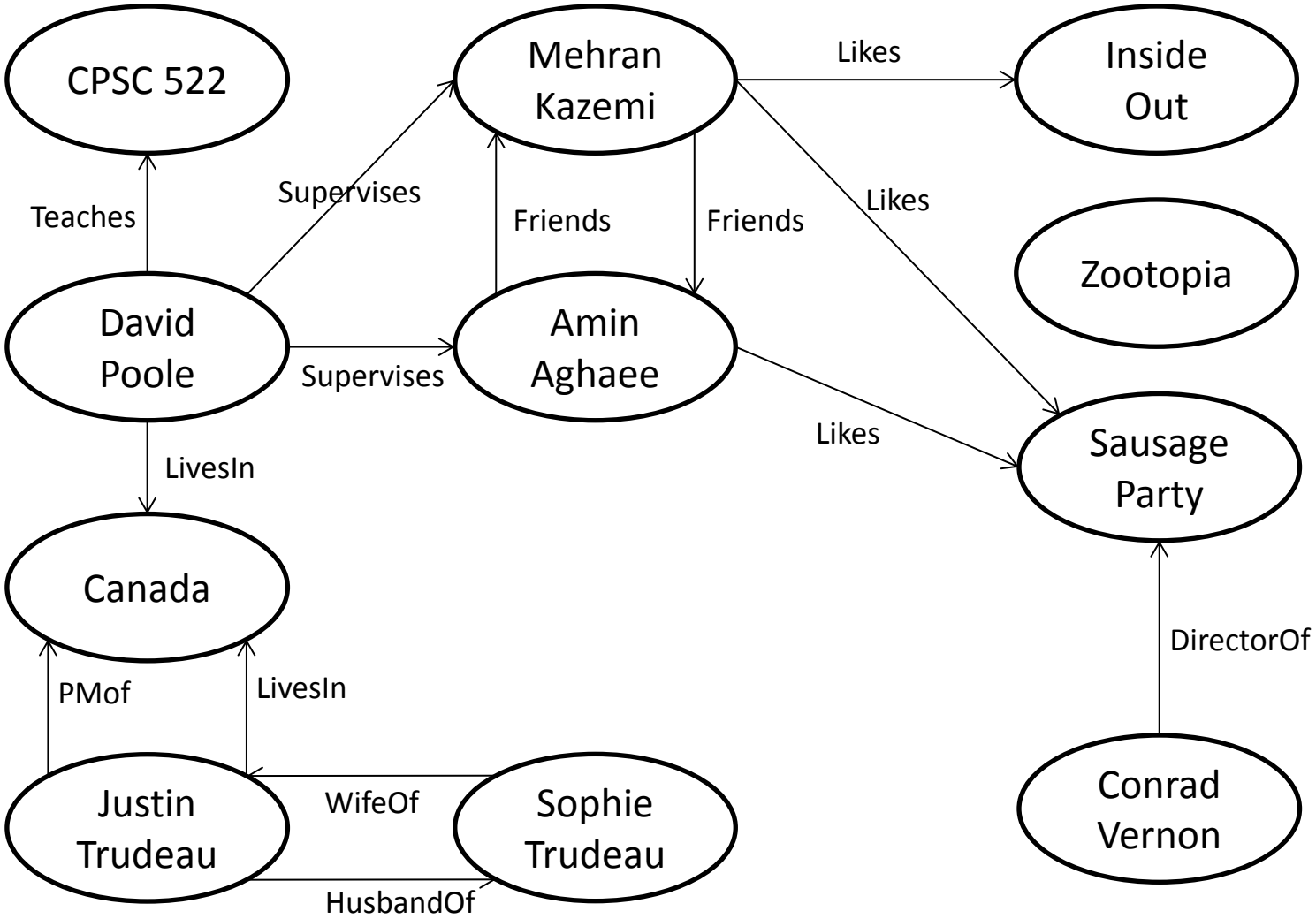


Alice:	1.2	−1.9	0.9	1.1	−0.9
Ben:	−1	−2	3	1	1.2
Chris:	3.1	−1	−0.2	0.2	0.3
GF:	1.5	0.2	−2	−9	1
Her:	0.1	0	−0.9	1.1	1.9

…

These vectors are called
Embeddings

What if we have more relations?



What if we have more relations?

Does Sophie Trudeau live in Canada?

Sophie Trudeau:

v_1	v_2	v_3	v_4	v_5

Canada:

w_1	w_2	w_3	w_4	w_5

What if we have more relations?

Does Sophie Trudeau live in Canada?

Sophie Trudeau:	v_1	v_2	v_3	v_4	v_5
LivesIn:	r_1	r_2	r_3	r_4	r_5
Canada:	w_1	w_2	w_3	w_4	w_5

What if we have more relations?

Does Sophie Trudeau live in Canada?

Sophie Trudeau:	v_1	v_2	v_3	v_4	v_5
LivesIn:	r_1	r_2	r_3	r_4	r_5
Canada:	w_1	w_2	w_3	w_4	w_5

$$v_1 r_1 w_1 + v_2 r_2 w_2 + v_3 r_3 w_3 + v_4 r_4 w_4 + v_5 r_5 w_5$$

What if we have more relations?

Does Sophie Trudeau live in Canada?

Sophie Trudeau:	v_1	v_2	v_3	v_4	v_5
LivesIn:	r_1	r_2	r_3	r_4	r_5
Canada:	w_1	w_2	w_3	w_4	w_5

$$v_1 r_1 w_1 + v_2 r_2 w_2 + v_3 r_3 w_3 + v_4 r_4 w_4 + v_5 r_5 w_5$$

This approach is called ***DistMult*** *(ICLR-2015)*

DistMult Enforces Symmetry

Does David supervise Mehran?

David:	d_1	d_2	d	d_4	d_5
Supervises:	s_1	s_2	s_3	s_4	s_5
Mehran:	m_1	m_2	m_3	m_4	m_5

$$d_1 s_1 m_1 + d_2 s_2 m_2 + d_3 s_3 m_3 + d_4 s_4 m_4 + d_5 s_5 m_5$$

Do you see the problem?

DistMult Enforces Symmetry

Does Mehran supervise David?

Mehran:	m_1	m_2	m_3	m_4	m_5
Supervises:	s_1	s_2	s_3	s_4	s_5
David:	d_1	d_2	d	d_4	d_5

$$d_1 s_1 m_1 + d_2 s_2 m_2 + d_3 s_3 m_3 + d_4 s_4 m_4 + d_5 s_5 m_5$$

ComplEx (ICML-2016)

David:	$d_1 + id'_1$	$d_2 + id'_2$	$d_3 + id'_3$	$d_4 + id'_4$	$d_5 + id'_5$
Supervises:	$s_1 + is'_1$	$s_2 + is'_2$	$s_3 + is'_3$	$s_4 + is'_4$	$s_5 + is'_5$
Mehran:	$m_1 + im'_1$	$m_2 + im'_2$	$m_3 + im'_3$	$m_4 + im'_4$	$m_5 + im'_5$

ComplEx (ICML-2016)

Does David supervise Mehran?

David:	$d_1 + id'_1$	$d_2 + id'_2$	$d_3 + id'_3$	$d_4 + id'_4$	$d_5 + id'_5$
Supervises:	$s_1 + is'_1$	$s_2 + is'_2$	$s_3 + is'_3$	$s_4 + is'_4$	$s_5 + is'_5$
Mehran:	$m_1 - im'_1$	$m_2 - im'_2$	$m_3 - im'_3$	$m_4 - im'_4$	$m_5 - im'_5$

$$d_1 s_1 m_1 - i\, d_1 s_1 m'_1 + i\, d_1 s'_1 m_1 + d_1 s'_1 m'_1 + i d'_1 s_1 m_1 + d'_1 s_1 m'_1 - d'_1 s'_1 m_1 + i d'_1 s'_1 m'_1 + \ \ldots$$

ComplEx (ICML-2016)

Does David supervise Mehran?

David:	$d_1 + id'_1$	$d_2 + id'_2$	$d_3 + id'_3$	$d_4 + id'_4$	$d_5 + id'_5$
Supervises:	$s_1 + is'_1$	$s_2 + is'_2$	$s_3 + is'_3$	$s_4 + is'_4$	$s_5 + is'_5$
Mehran:	$m_1 - im'_1$	$m_2 - im'_2$	$m_3 - im'_3$	$m_4 - im'_4$	$m_5 - im'_5$

$$Re(d_1 s_1 m_1 - i\, d_1 s_1 m'_1 + i\, d_1 s'_1 m_1 + d_1 s'_1 m'_1 + i d'_1 s_1 m_1 + d'_1 s_1 m'_1 - d'_1 s'_1 m_1 + i d'_1 s'_1 m'_1 + \ldots)$$

$$= d_1 s_1 m_1 + d_1 s'_1 m'_1 + d'_1 s_1 m'_1 - d'_1 s'_1 m_1 + \cdots$$

ComplEx (ICML-2016)

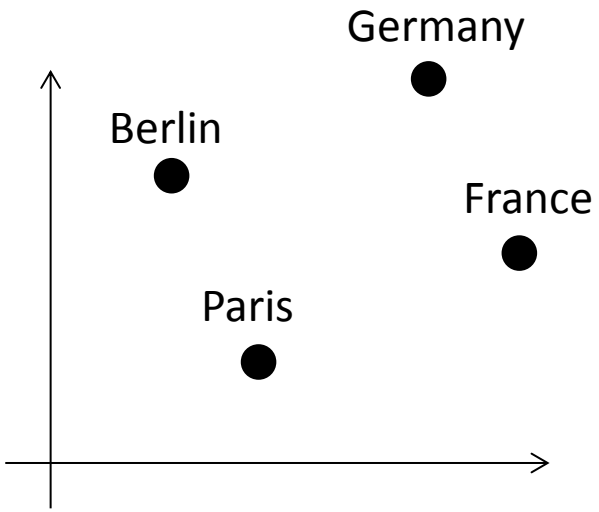
Does Mehran supervise David?

Mehran:	$m_1 + im'_1$	$m_2 + im'_2$	$m_3 + im'_3$	$m_4 + im'_4$	$m_5 + im'_5$
Supervises:	$s_1 + is'_1$	$s_2 + is'_2$	$s_3 + is'_3$	$s_4 + is'_4$	$s_5 + is'_5$
David:	$d_1 - id'_1$	$d_2 - id'_2$	$d_3 - id'_3$	$d_4 - id'_4$	$d_5 - id'_5$

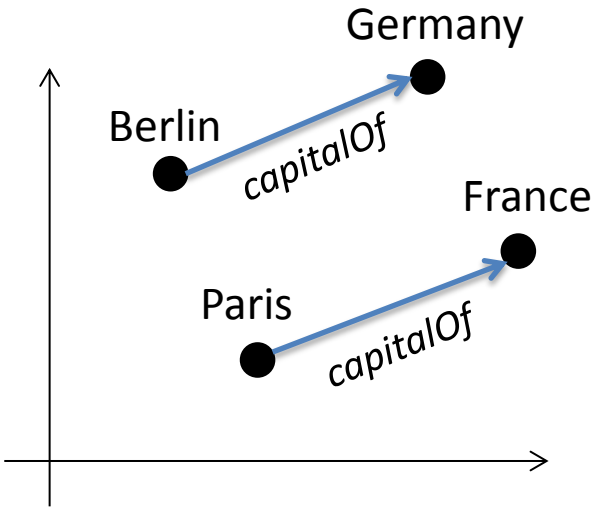
$$Re(d_1 s_1 m_1 + i\, d_1 s_1 m'_1 + i\, d_1 s'_1 m_1 - d_1 s'_1 m'_1 - i d'_1 s_1 m_1 + d'_1 s_1 m'_1 + d'_1 s'_1 m_1 + i d'_1 s'_1 m'_1 + \ \ldots)$$

$$= d_1 s_1 m_1 - d_1 s'_1 m'_1 + d'_1 s_1 m'_1 + d'_1 s'_1 m_1 + \cdots$$

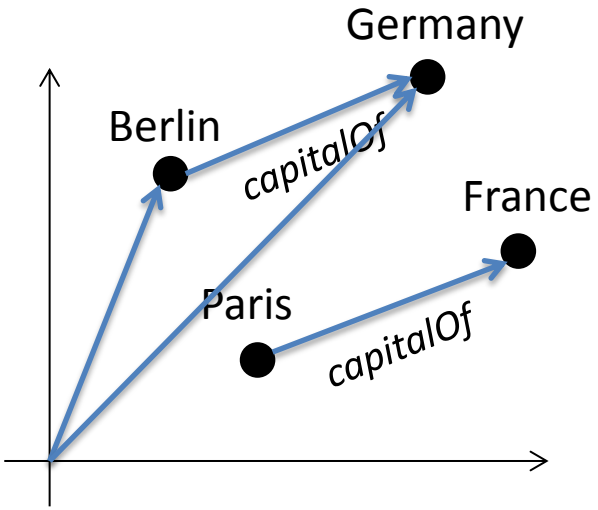
Translational Approaches



Translational Approaches



Translational Approaches



$$V(B) + V(cOf) \approx V(G)$$

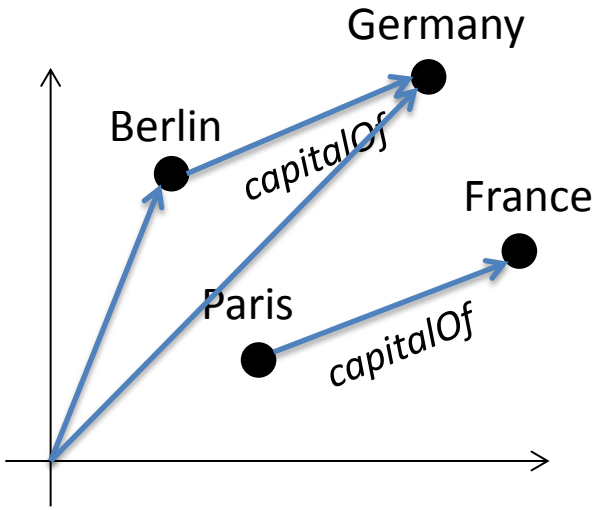
Or equivalently

$$||V(B) + V(cOf) - V(G)|| \approx 0$$

TransE (NIPS-2013)

Sophie Trudeau:	v_1	v_2	v_3	v_4	v_5
LivesIn:	r_1	r_2	r_3	r_4	r_5
Canada:	w_1	w_2	w_3	w_4	w_5

Whether ST lives in Canada is conversely correlated with $||V + R - W||$

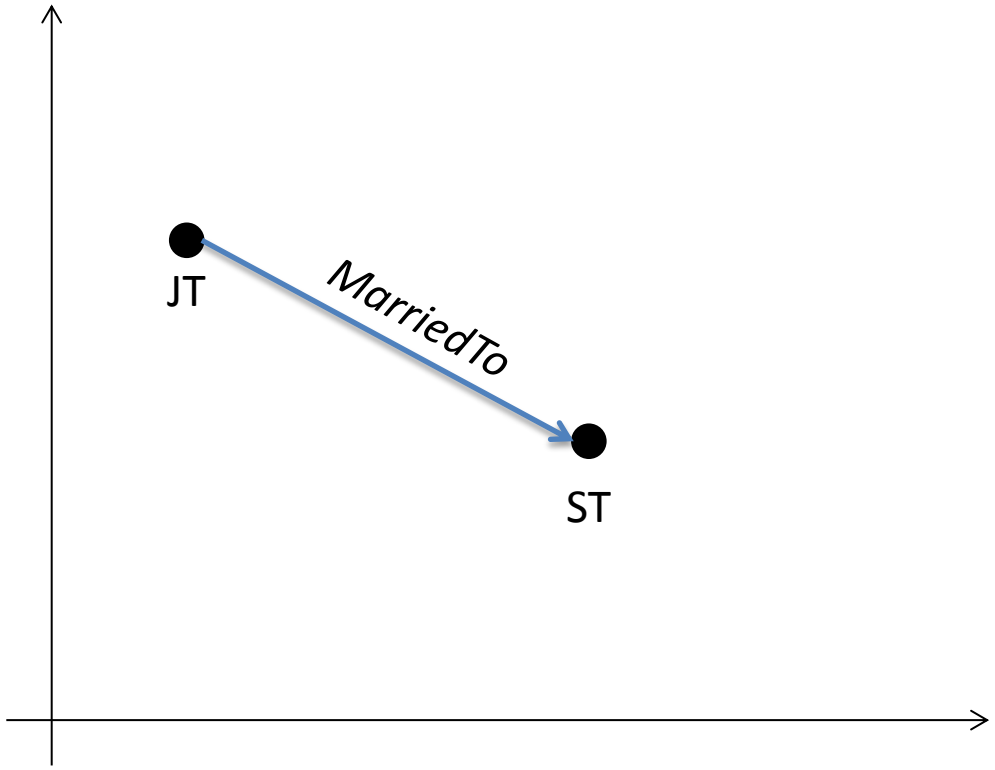


$$V(B) + V(cOf) \approx V(G)$$

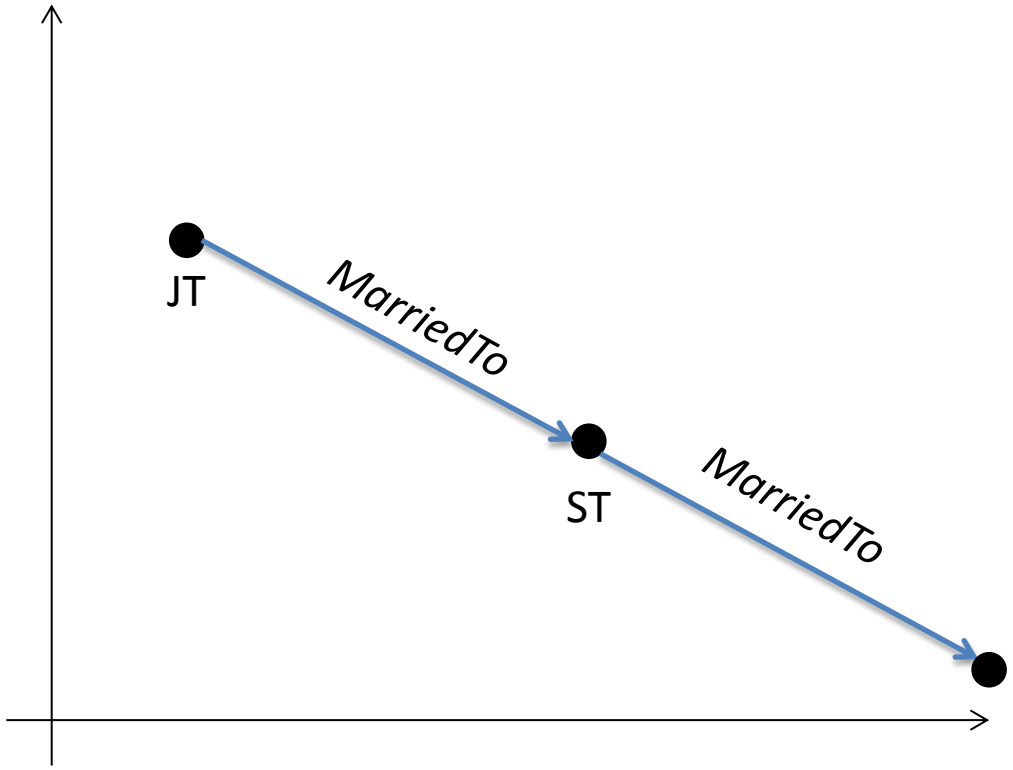
Or equivalently

$$||V(B) + V(cOf) - V(G)|| \approx 0$$

TransE and Symmetric Relations

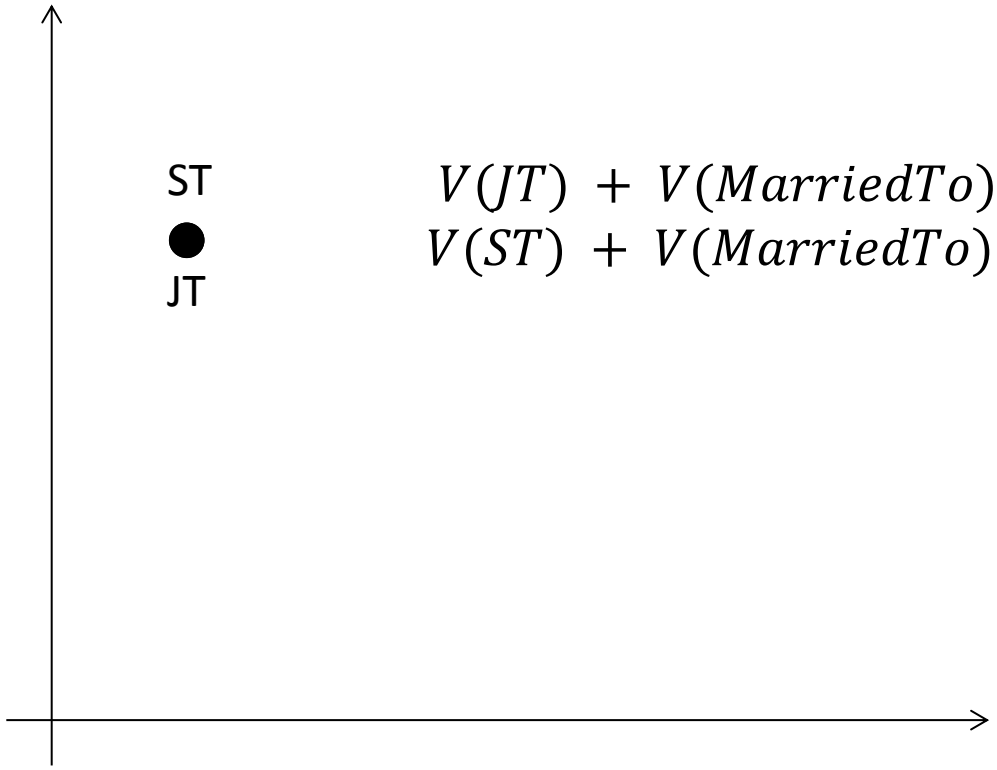


TransE and Symmetric Relations



TransE and Symmetric Relations

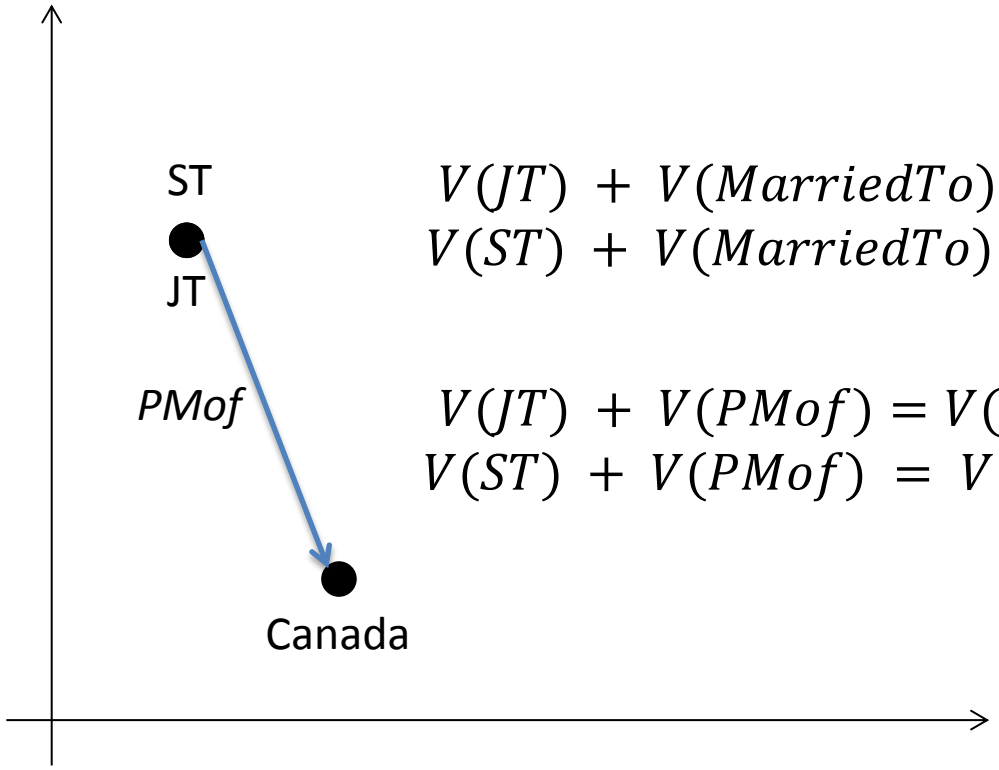
What if R is a vector of zeros?



$$V(JT) + V(MarriedTo) = V(ST)$$
$$V(ST) + V(MarriedTo) = V(JT)$$

TransE and Symmetric Relations

What if R is a vector of zeros?



$$V(JT) + V(MarriedTo) = V(ST)$$
$$V(ST) + V(MarriedTo) = V(JT)$$

$$V(JT) + V(PMof) = V(Canada)$$
$$V(ST) + V(PMof) = V(Canada)$$

Variants of TransE

Many variants of TransE have been proposed to address the issues in TransE:

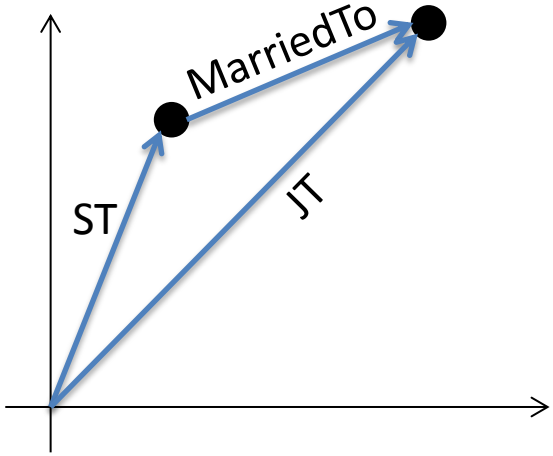
- TransH
- TransR
- CTransR
- TransD
- TransG
- STransE
- FTransE
- PTransE
- RTransE

STransE (ACL-HLT-2016)

Sophie Trudeau: V(ST)

MarriedTo: V(MT)

Justin Trudeau: V(JT)



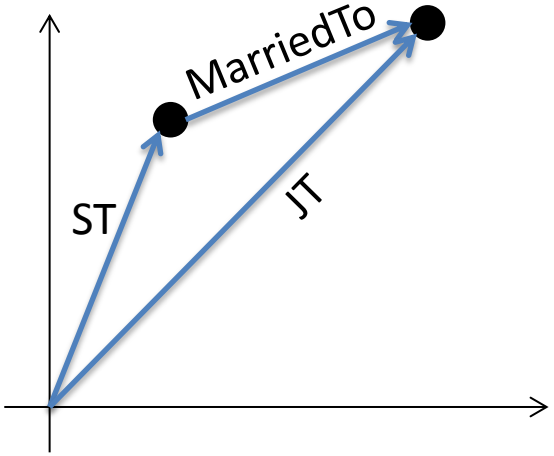
TransE: $V(ST) + V(MT) \approx V(JT)$

STransE (ACL-HLT-2016)

Sophie Trudeau: V(ST)

MarriedTo: V(MT), M1(MT), M2(MT)

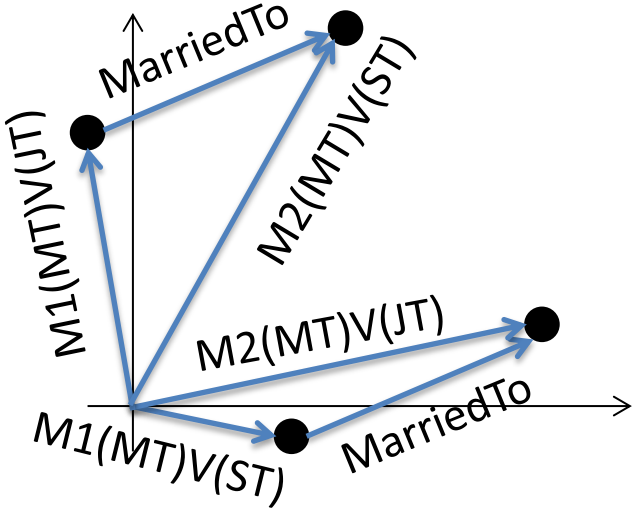
Justin Trudeau: V(JT)



TransE: $V(ST) + V(MT) \approx V(JT)$

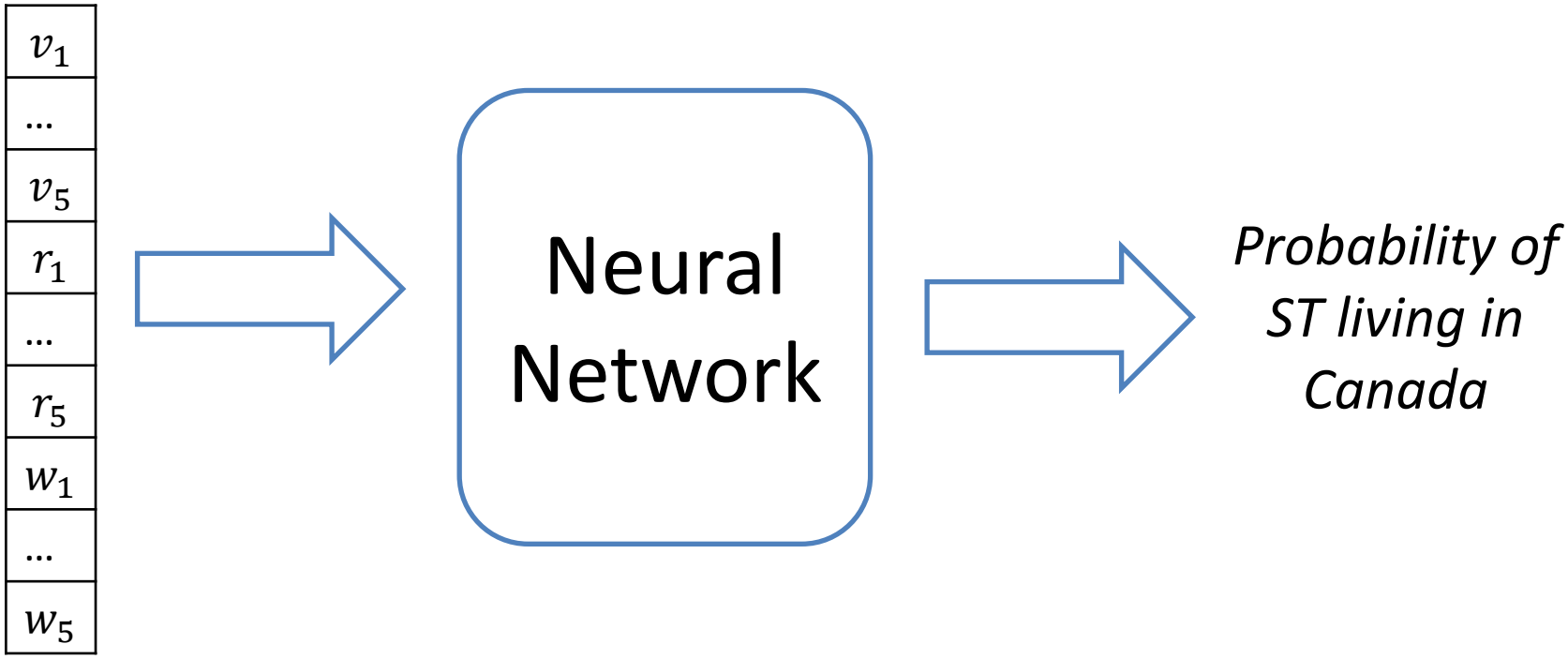
STransE: $\mathrm{M1(MT)}V(ST) + V(MT) \approx M2(MT)V(JT)$

$\mathrm{M1(MT)}V(JT) + V(MT) \approx M2(MT)V(ST)$

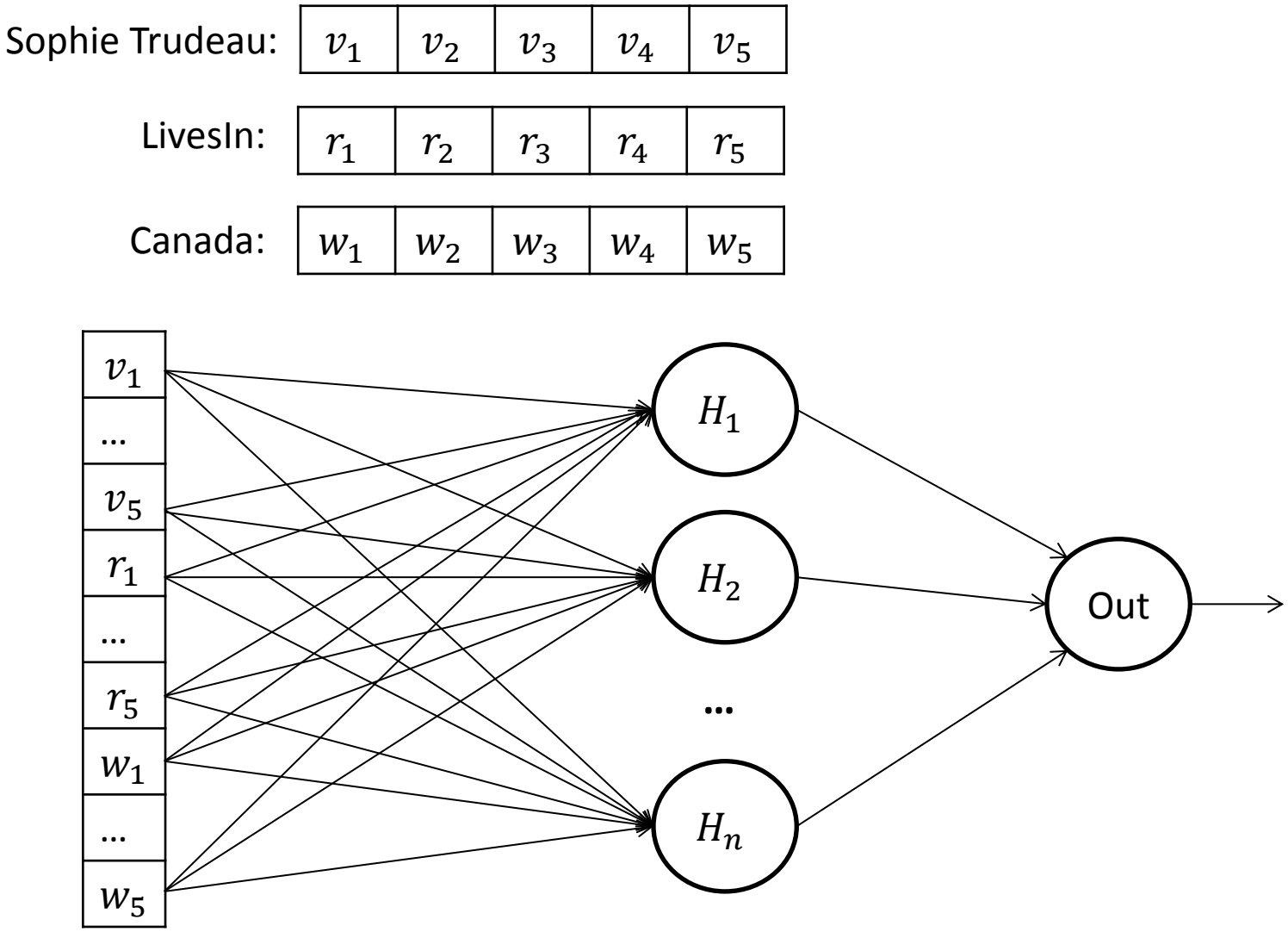


Deep Learning

Sophie Trudeau:	v_1	v_2	v_3	v_4	v_5
LivesIn:	r_1	r_2	r_3	r_4	r_5
Canada:	w_1	w_2	w_3	w_4	w_5

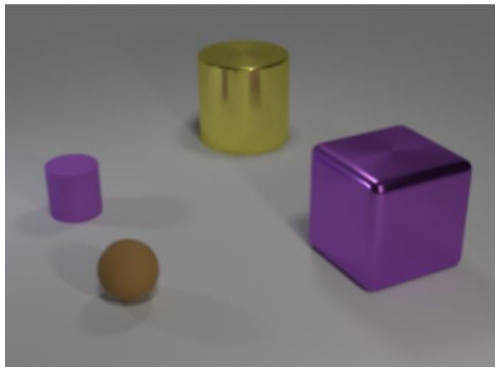


ER-MLP (SIGKDD-2014)



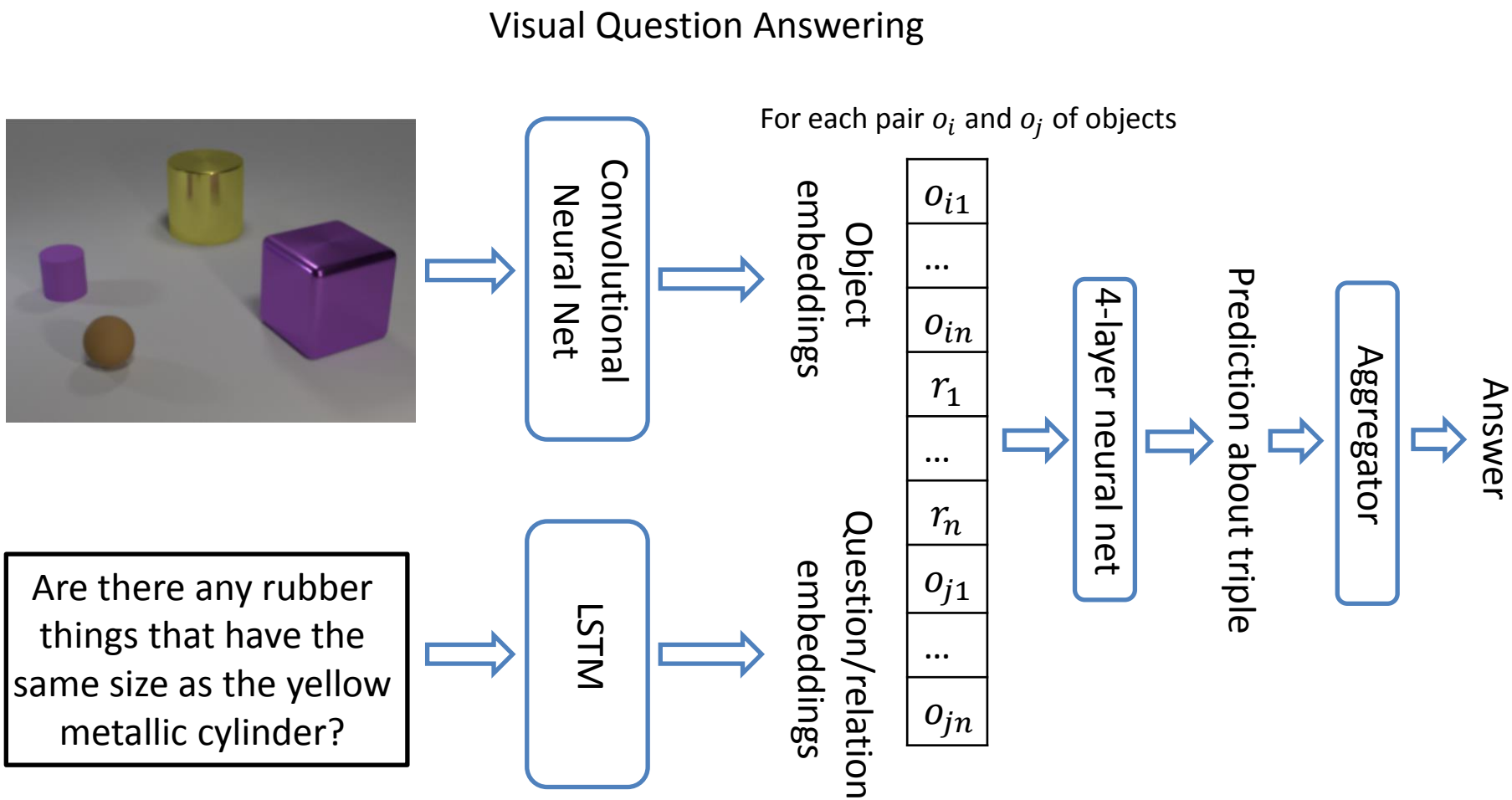
Sandop et al. (NIPS-2017)

Visual Question Answering



Are there any rubber things that have the same size as the yellow metallic cylinder?

Sandop et al. (NIPS-2017)



Comparisons

	Filtered MRR on WN18	Filtered MRR on FB15k
DistMult		
ComplEx		
TransE		
STransE		
ER-MLP		

- DistMult: Sum of the element-wise product of three real-valued vectors
- ComplEx: Sum of the element-wise product of three complex-valued vectors
- TransE: Sum of the head entity and relation vectors gives the tail entity vector
- STransE: Project the entities into a relation-specific space and apply TransE
- ER-MLP: Concatenate the embeddings and feed it into a neural network

Comparisons

	Filtered MRR on WN18	Filtered MRR on FB15k
DistMult	0.822	0.654
ComplEx	0.941	0.692
TransE	0.454	0.380
STransE	0.657	0.543
ER-MLP	0.712	0.288

- DistMult: Sum of the element-wise product of three real-valued vectors
- ComplEx: Sum of the element-wise product of three complex-valued vectors
- TransE: Sum of the head entity and relation vectors gives the tail entity vector
- STransE: Project the entities into a relation-specific space and apply TransE
- ER-MLP: Concatenate the embeddings and feed it into a neural network

If you want to do your course project on knowledge graph completion, come talk to me (room X568), or send me an email at ***smkazemi@cs.ubc.ca***