Never-Ending Language Learning

Tom Mitchell, William Cohen, and Many Collaborators
Carnegie Mellon University

We will never really understand learning until we build machines that

- learn many different things,
- from years of diverse experience,
- in a staged, curricular fashion,
- and become better <u>learners</u> over time.

Tenet 2:

Natural language understanding requires a belief system

A natural language understanding system should react to text by saying either:

- I understand, and already knew that
- I understand, and didn't know, but accept it
- I understand, and disagree because …

NELL: Never-Ending Language Learner

Inputs:

- initial ontology (categories and relations)
- dozen examples of each ontology predicate
- the web
- occasional interaction with human trainers

The task:

- run 24x7, forever
- each day:
    1. extract more facts from the web to populate the ontology
    2. learn to read (perform #1) better than yesterday

NELL today

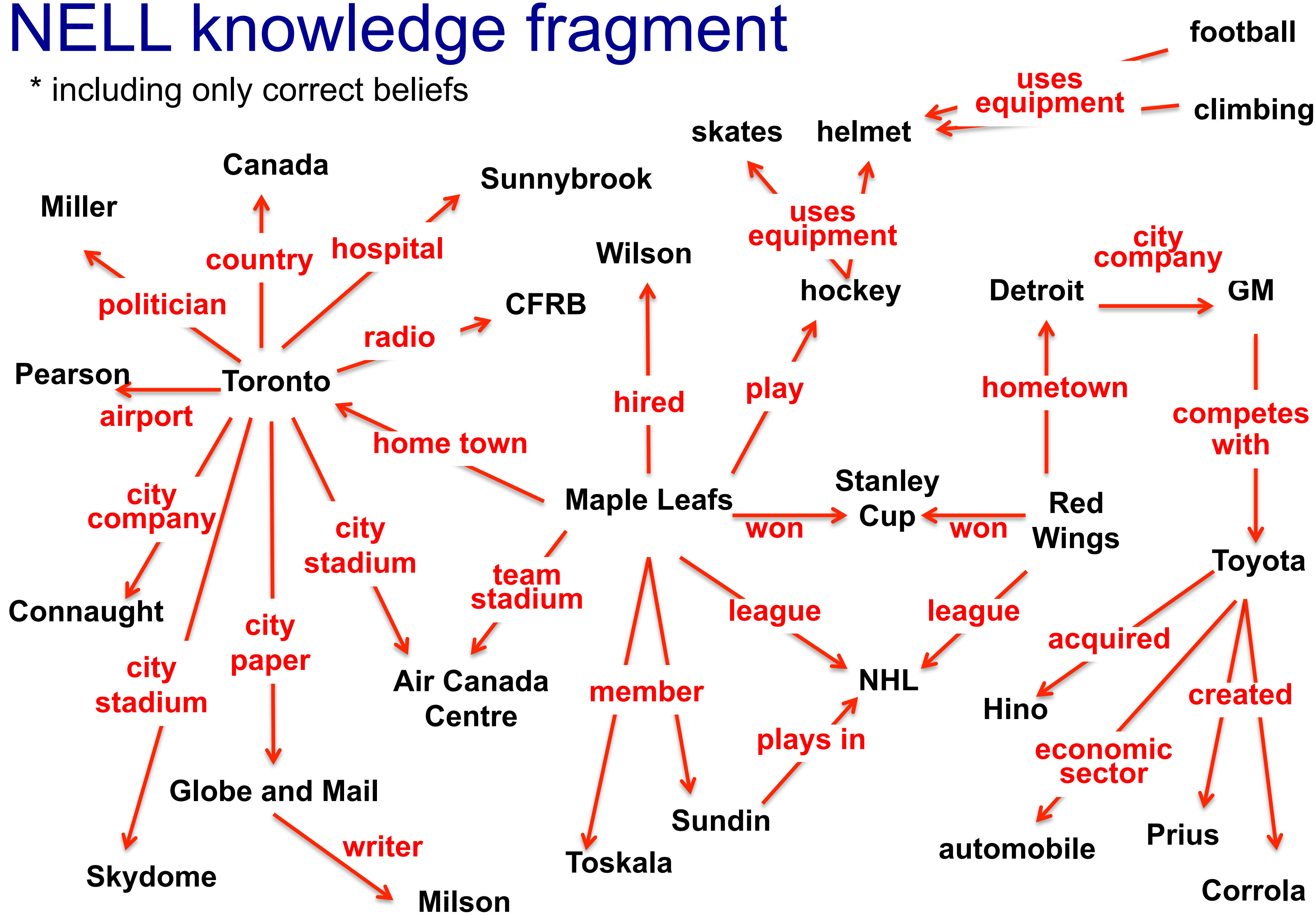
Running 24x7, since January, 12, 2010

Result:

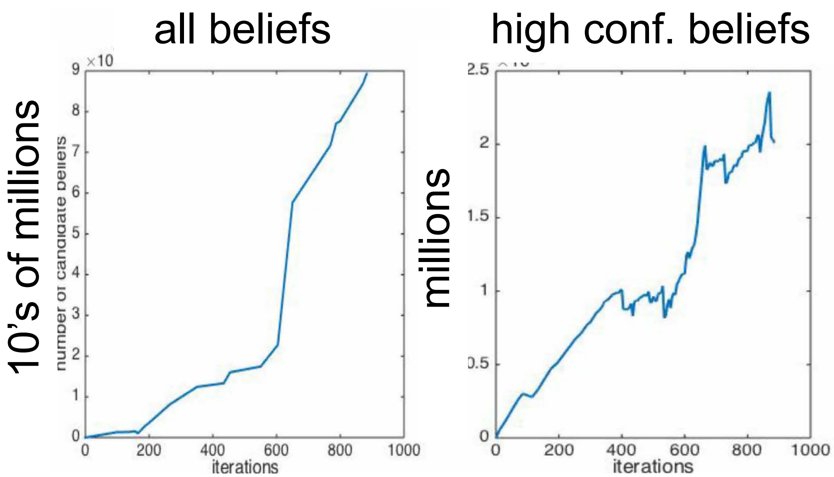
- knowledge base with 90 million candidate beliefs
- learning to read
- learning to reason
- extending ontology

NELL knowledge fragment

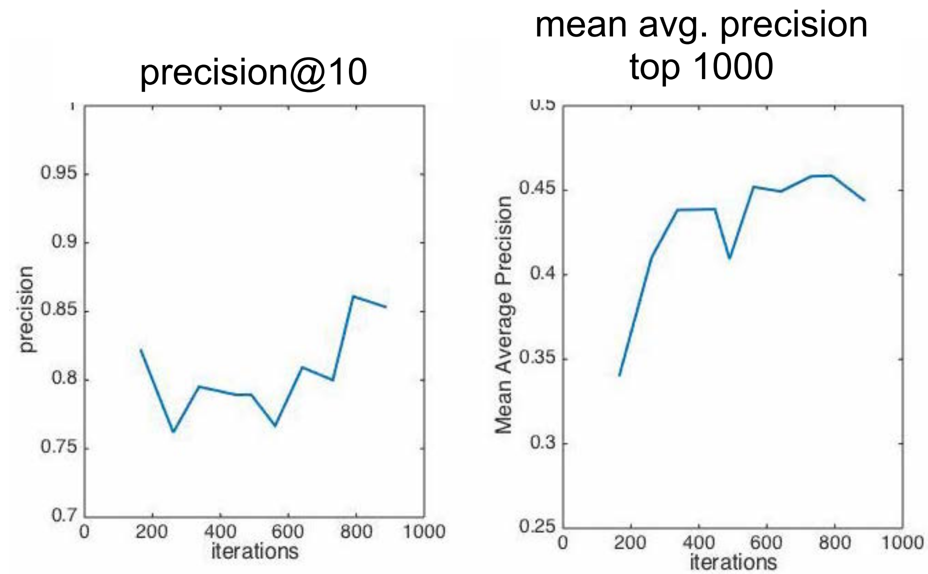
* including only correct beliefs



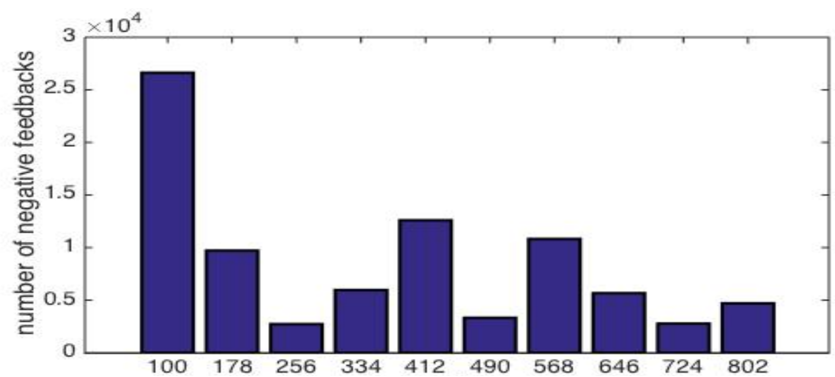
NELL Is Improving Over Time (Jan 2010 to Nov 2014)



number of NELL beliefs vs. time



reading accuracy vs. time
(average over 31 predicates)



human feedback vs. time
(average 2.4 feedbacks per predicate per month)

NELL Today

- eg. “diabetes”, “Avandia”, “tea”, “IBM”, “love” “baseball” “San Juan” “BacteriaCausesCondition” “kitchenItem” “ClothingGoesWithClothing” …

Recently-Learned Facts

instance	iteration	date learned
jerry_patton is a chef	881	24-oct-2014
flavones is a video game	883	02-nov-2014
smith_s_rose_bellied_lizard is a reptile	883	02-nov-2014
gray_flycatcher is a bird	883	02-nov-2014
basalt_plains can be a part of a landscape	883	02-nov-2014
louisiana is a state or province located in the geopolitical location u_s_	886	21-nov-2014
kansas_city_chiefs is a sports team that won the super_bowl	886	21-nov-2014
miller has been charged with contempt	886	21-nov-2014
the companies new_york and london_sunday_times compete with eachother	884	08-nov-2014
aberdeen is a city located in the geopolitical location the_united_kingdom	886	21-nov-2014

Portuguese NELL

[Estevam Hruschka, 2014]

Recent

instance

adriane_g
basf_e_fa
manaus_
jacutinga
fim_da_g
bamerind
nissan is
susana_v
campeon
toyota_m

- mes
- ano
- dataliteral
- evento
  - eventoesportista
    - olimpiadas
    - grandepremio
    - corrida
    - jogoesportivo
  - convencao
  - fenomenometeo
  - tipodeeventomil
    - conflitomilitar
  - conferencia
    - conferenciade
  - eleicao
  - festivaldemusica
  - festivaldefilmes
  - resultadodeeven
  - crimeoucusaca
- contapolitica
- coordernadas
- metricadeam
- emocao

conflitomilitar

(category)

See learned instances of conflitomilitar as a list or on a map

Metadata

- **allLearnedPatterns**
  - "a armada durante _" "a causa diplom�tica _" "a ci on _" "a armamentista durante _" "a declara��o de capitula��o _ data _" "a disputa teconol�gica _" "a fronteira interalem� imin�ncia _" "a guarni��o francesa durante _" "a indepe _" "a P.Y.S.B.E. na _" "a P.Y.S.B.E. _" "a ponte resqu�cios _ promover� _" "acabaram a produ��o no _" "acusa��o d "agudiza��o no _" "antecederam os conflitos da _" "antia "arquiinimigos na _" "As d�cadas do cabar� Ap�s _" "As _" "As origens do conflito A _" "as raz�es te�ricas para _" iraquianas durante _" "avi�es de luta e _" "bacilos durante "batalha da propaganda durante _" "bimotor na _" "catrefa _" "cidades do leste durante _" "combates de avi�o _" "co _" "conflito militar apelidado de _" "conflito militar assim c militar chamado de _" "conflito militar como _" "conflito mi _" "conflito militar tal como _" "conflitos militares assim co

How does NELL work?

Semi-Supervised Bootstrap Learning

it's underconstrained!!

Learn which noun phrases are cities:

Paris
Pittsburgh
Seattle
Montpelier

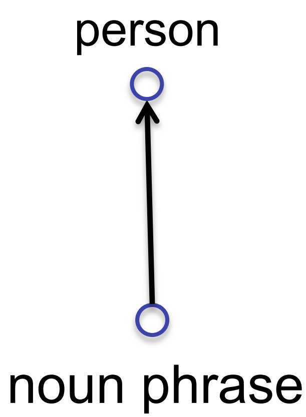
San Francisco
Berlin
denial

anxiety
selfishness
London

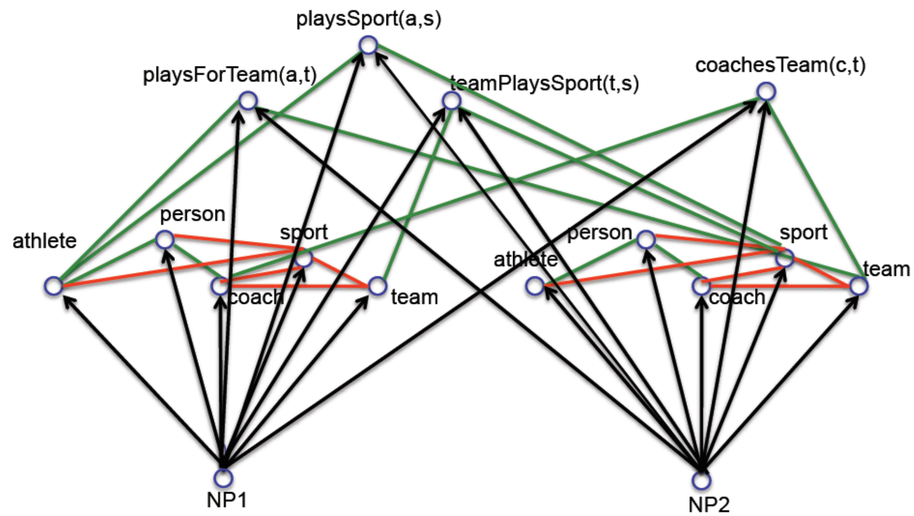
mayor of arg1
live in arg1

arg1 is home of
traits such as arg1

Key Idea 1: Coupled semi-supervised training of many functions



hard
(underconstrained)
semi-supervised
learning problem

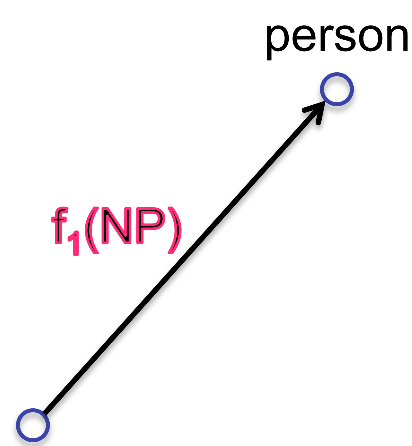


much easier (more constrained)
semi-supervised learning problem

Type 1 Coupling: Co-Training, Multi-View Learning

Supervised training of 1 function:

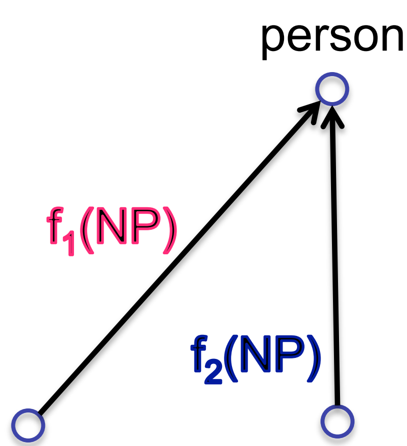
Minimize: $$\sum_{<np,person> \in labeled\ data} |f_1(np) - person|$$



NP: NP context distribution

__ is a friend
rang the __
…
__ walked in

Type 1 Coupling: Co-Training, Multi-View Learning



NP: NP context distribution | NP morphology

__ is a friend
rang the __
...
__ walked in

capitalized?
ends with '...ski'?
...
contains "univ."?

Coupled training of 2 functions:

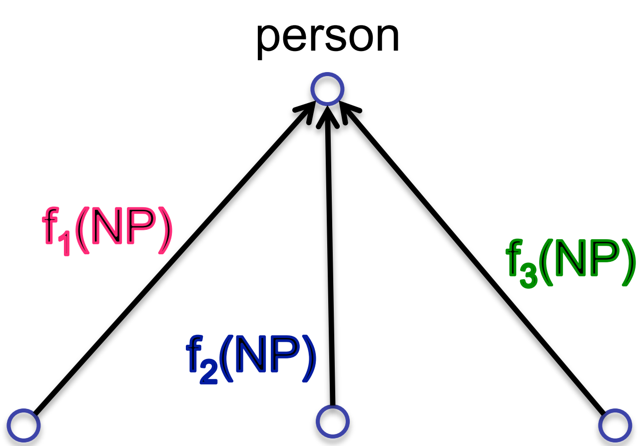
Minimize: $$\sum_{<np,person> \,\in\, labeled\ data} |f_1(np) - person|$$

$$+ \sum_{<np,person> \,\in\, labeled\ data} |f_2(np) - person|$$

$$+ \sum_{np \,\in\, unlabeled\ data} |f_1(np) - f_2(np)|$$

Type 1 Coupling: Co-Training, Multi-View Learning

[Blum & Mitchell; 98]
[Dasgupta et al; 01]
[Ganchev et al., 08]
[Sridharan & Kakade, 08]
[Wang & Zhou, ICML10]



NP:

NP context distribution	NP morphology	NP HTML contexts
__ is a friend	*capitalized?*	*www.celebrities.com:*
rang the __	*ends with '...ski'?*	* __ *
...	...	
__ walked in	*contains "univ."?*	...

NELL: Learned reading strategies

Mountain:

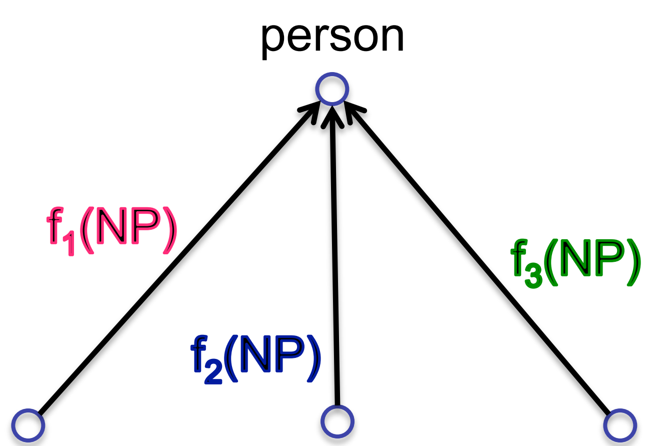
"volcanic crater of _" "volcanic erupt
region of _" "volcano , called _" "vo
"volcano known as _" "volcano Mt _
including _" "volcanoes , like _" "vo
_" "volcanoes including _" "volcano
"weather atop _" "weather station at
through _" "West face of _" "West r
ledge in _" "white summit of _" "wh
surrounding _" "wilderness areas ar
"winter ascents in _" "winter ascents
foothills of _" "world famous view of
popping by _" "you 've just climbed _
"_ ' crater" "_ ' eruption" "_ ' foothills
Camp" "_ 's drug guide" "_ 's east r
Face" "_ 's North Peak" "_ 's North
southeast ridge" "_ 's summit calder
's west ridge" "_ (D,DDD ft" " "_ clin
consult el diablo" "_ cooking planks"

Predicate	Feature	Weight
mountain	LAST=peak	1.791
mountain	LAST=mountain	1.093
mountain	FIRST=mountain	-0.875
musicArtist	LAST=band	1.853
musicArtist	POS=DT_NNS	1.412
musicArtist	POS=DT_JJ_NN	-0.807
newspaper	LAST=sun	1.330
newspaper	LAST=university	-0.318
newspaper	POS=NN_NNS	-0.798
university	LAST=college	2.076
university	PREFIX=uc	1.999
university	LAST=state	1.992
university	LAST=university	1.745
university	FIRST=college	-1.381
visualArtMovement	SUFFIX=ism	1.282
visualArtMovement	PREFIX=journ	-0.234

Predicate	Web URL	Extraction Template
academicField	http://scholendow.ais.msu.edu/student/ScholSearch.Asp	` [X] -`
athlete	http://www.quotes-search.com/d_occupation.aspx?o=+athlete	`<a href='d_author.aspx?a=[X]'>-`
bird	http://www.michaelforsberg.com/stock.html	`<option>[X]</option>`
bookAuthor	http://lifebehindthecurve.com/	`</li> <li>[X] by [Y] –`

Type 1 Coupling: Co-Training, Multi-View Learning

[Blum & Mitchell; 98]
[Dasgupta et al; 01]
[Ganchev et al., 08]
[Sridharan & Kakade, 08]
[Wang & Zhou, ICML10]

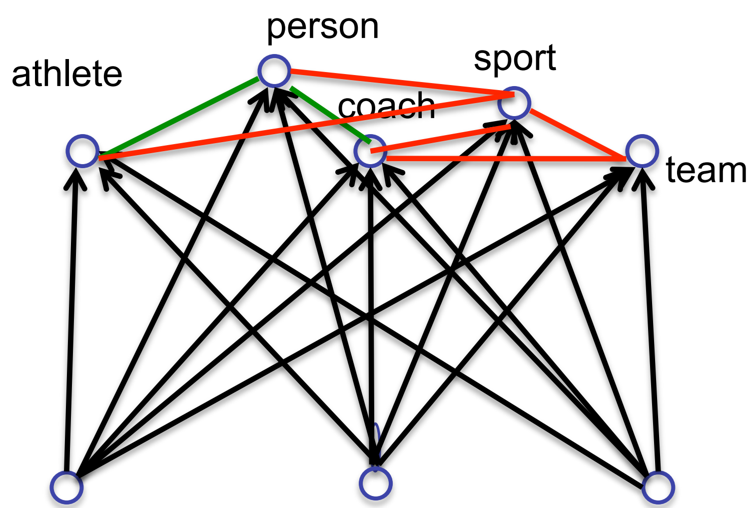


NP: NP context distribution | NP morphology | NP HTML contexts

NP context distribution	NP morphology	NP HTML contexts
__ is a friend	*capitalized?*	*www.celebrities.com:*
rang the __	*ends with '...ski'?*	* __ *
...	*...*	*...*
__ walked in	*contains "univ."?*	

Multi-view, Multi-Task Coupling

[Blum & Mitchell; 98]
[Dasgupta et al; 01]
[Ganchev et al., 08]
[Sridharan & Kakade, 08]
[Wang & Zhou, ICML10]
[Taskar et al., 2009]
[Carlson et al., 2009]



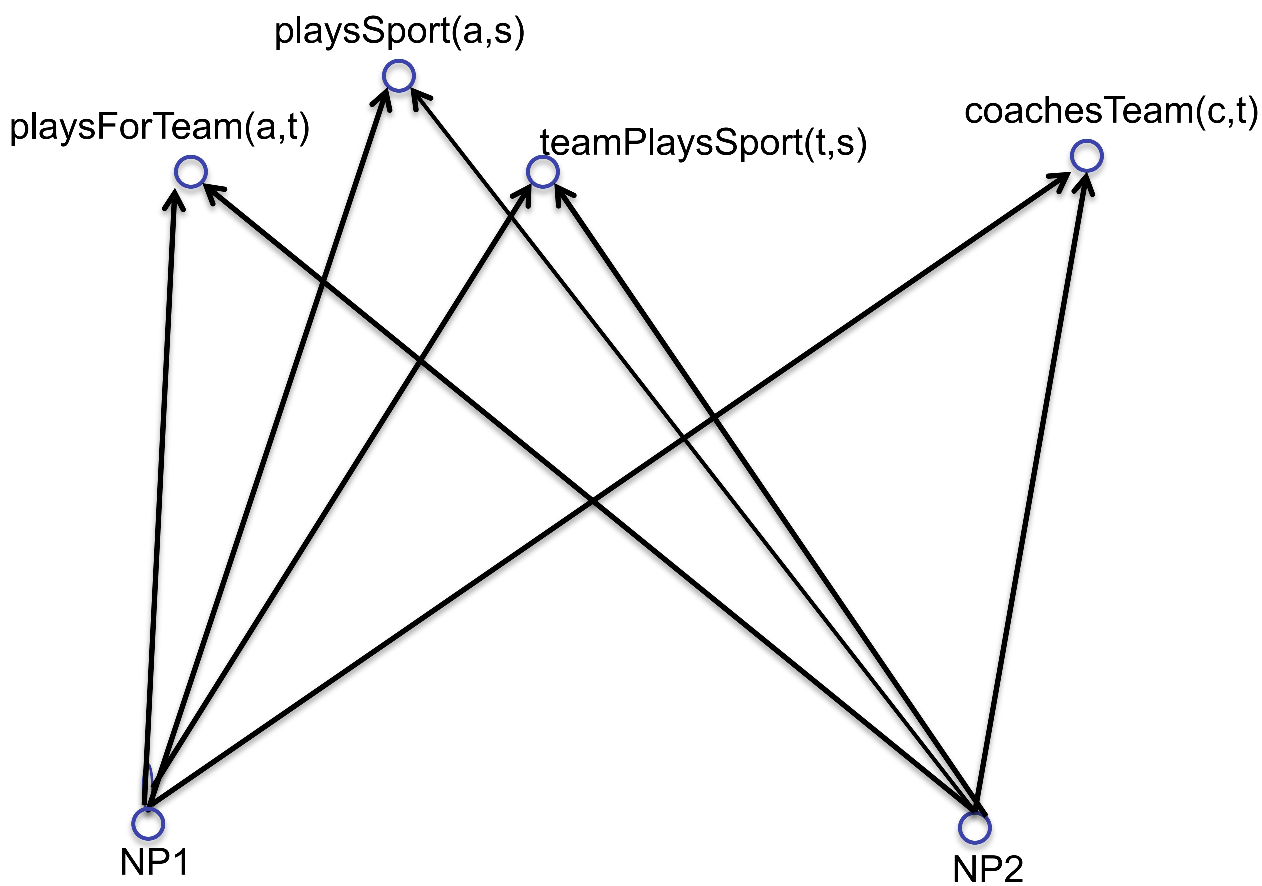
NP: NP text context distribution — NP morphology — NP HTML contexts

athlete(NP) → person(NP)

athlete(NP) → NOT sport(NP)

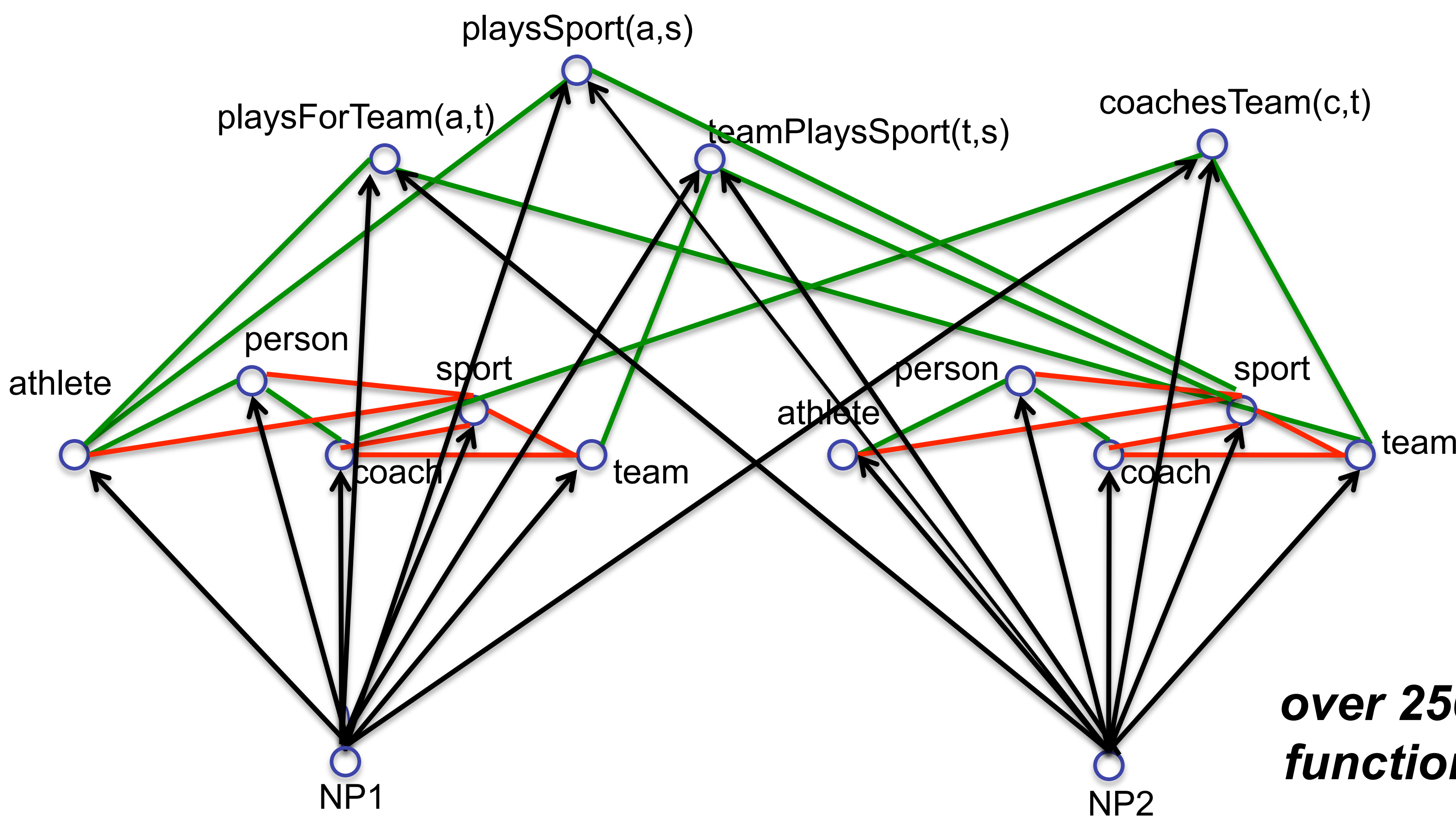
NOT athlete(NP) ← sport(NP)

Type 3 Coupling: Relation Argument Types



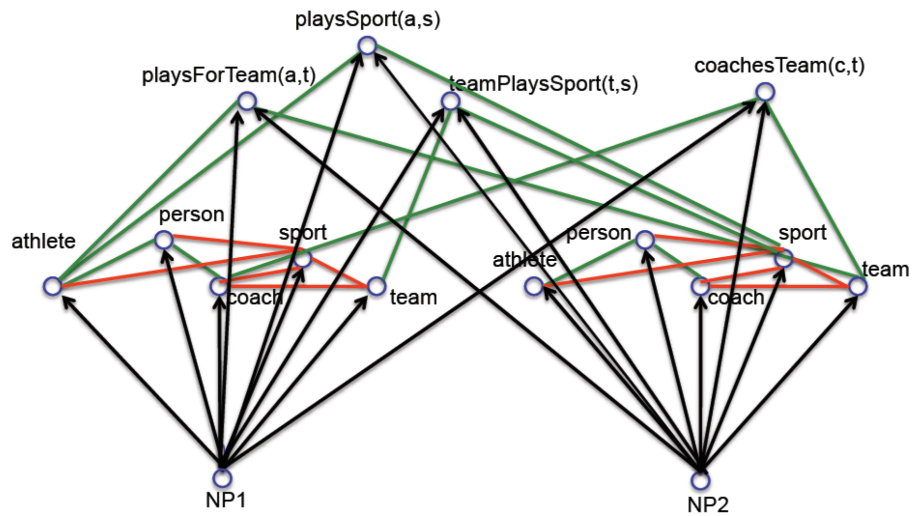
Type 3 Coupling: Relation Argument Types

playsSport(NP1,NP2) → athlete(NP1), sport(NP2)



over 2500 coupled functions in NELL

Pure EM Approach to Coupled Training



E: estimate labels for each function of each unlabeled example

M: retrain all functions, using these probabilistic labels

Scaling problem:

- **E** step: 25M NP's, 10^{14} NP pairs to label
- **M** step: 50M text contexts to consider for each function → 10^{10} parameters to retrain
- even more URL-HTML contexts…

NELL's Approximation to EM

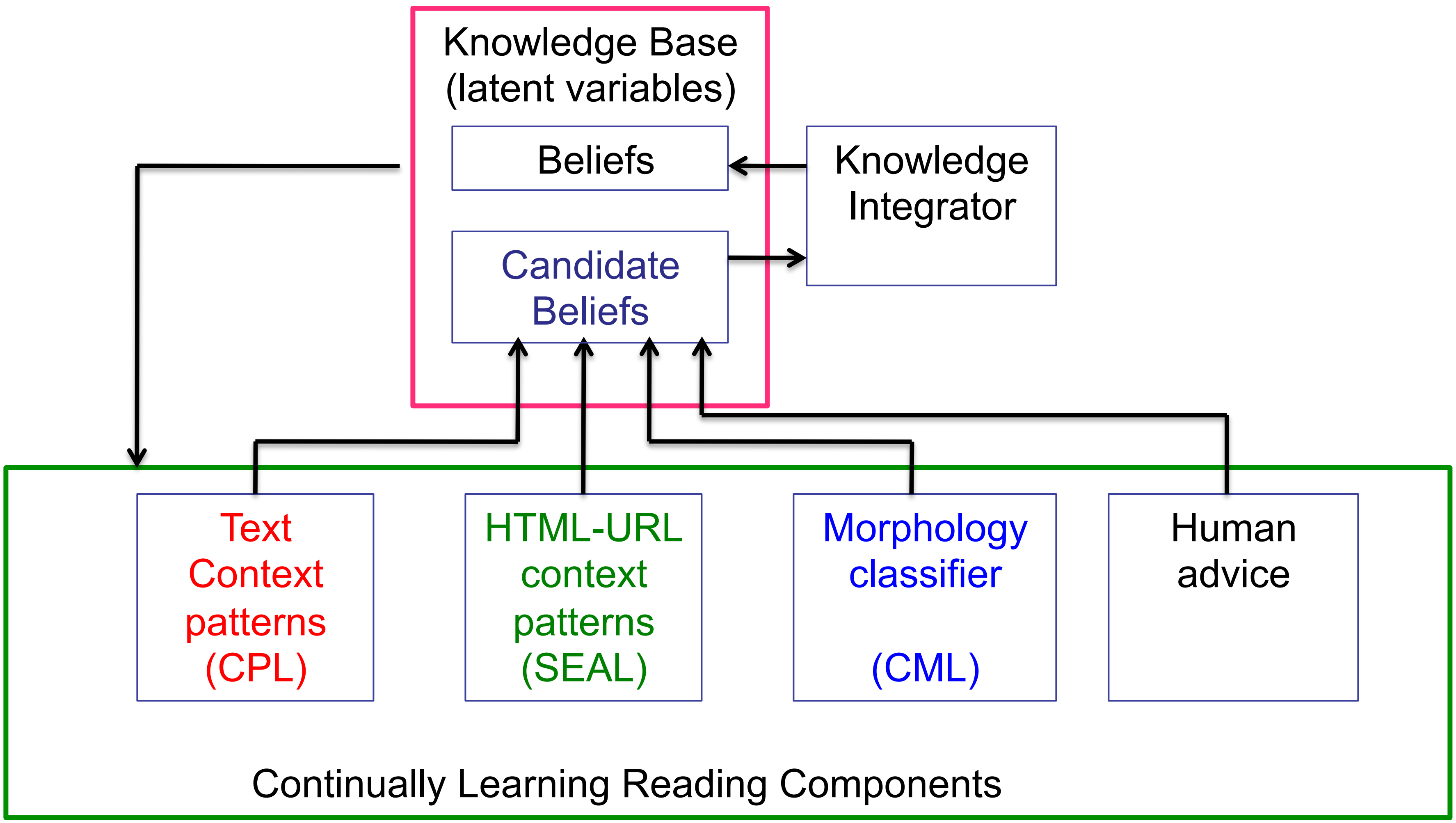
E' step:

- Re-estimate the knowledge base:
  - but consider only a growing subset of the latent variable assignments
  - category variables: up to 250 new NP's per category per iteration
  - relation variables: add only if confident and args of correct type
  - <u>this set of explicit latent assignments *</u>**<u>IS</u>**<u>* the knowledge base</u>

M' step:

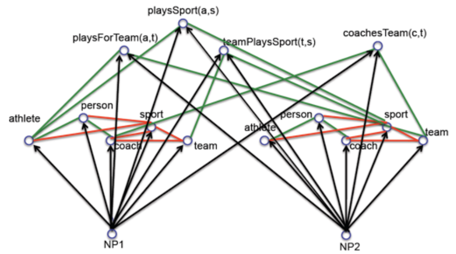
- Each view-based learner retrains itself from the updated KB
- "context" methods create growing subsets of contexts

Initial NELL Architecture



If coupled learning is the key,
how can we get new coupling constraints?

Key Idea 2:



Discover New Coupling Constraints

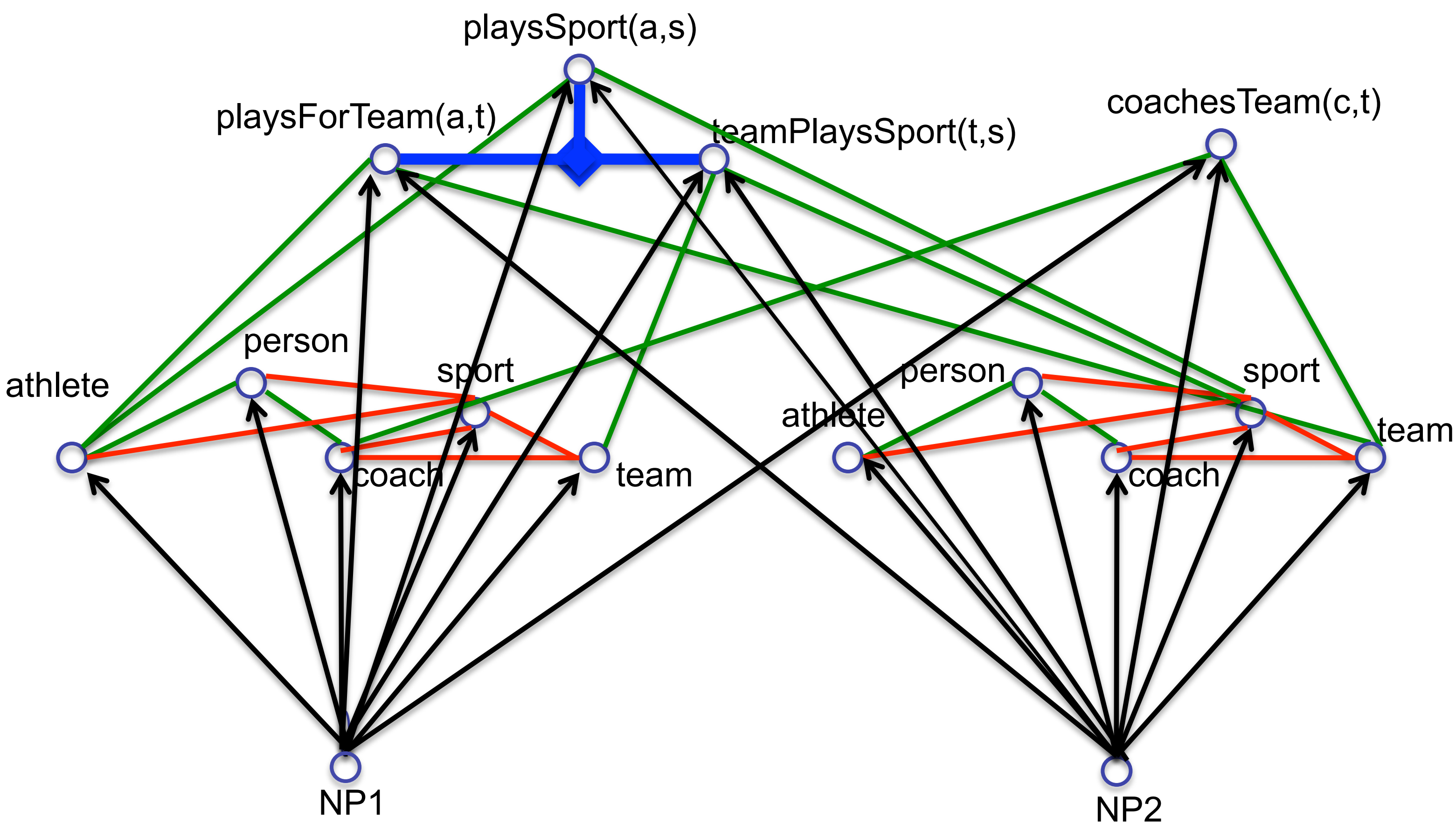
- learn horn clause rules/constraints:

0.93 athletePlaysSport(?x,?y) ← athletePlaysForTeam(?x,?z)
teamPlaysSport(?z,?y)

  - learned by data mining the knowledge base
  - connect previously uncoupled relation predicates
  - infer new unread beliefs
  - modified version of FOIL [Quinlan]

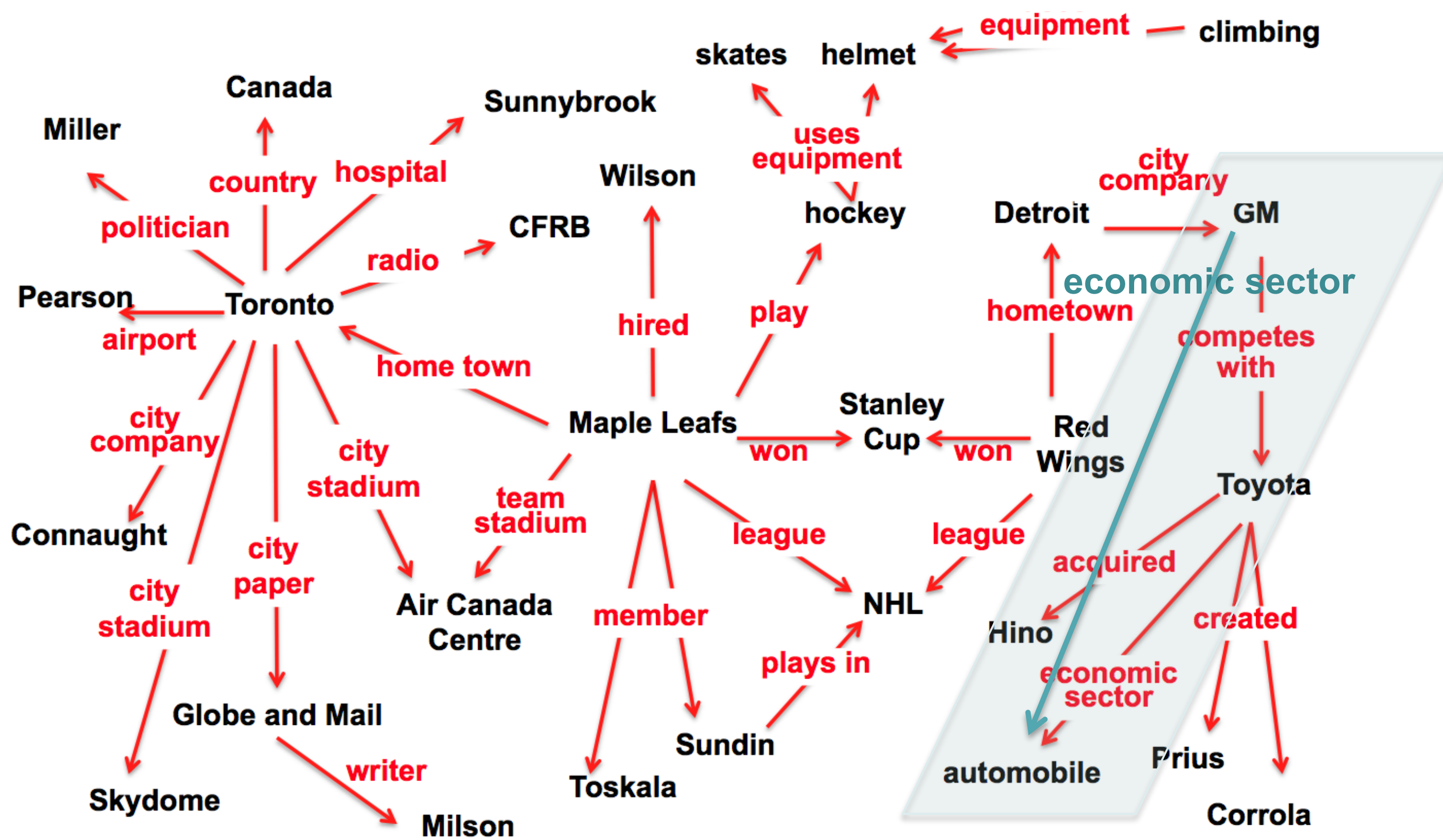
Learned Probabilistic Horn Clause Rules

0.93 playsSport(?x,?y) ← playsForTeam(?x,?z), teamPlaysSport(?z,?y)



Infer New Beliefs

[Lao, Mitchell, Cohen, *EMNLP* 2011]

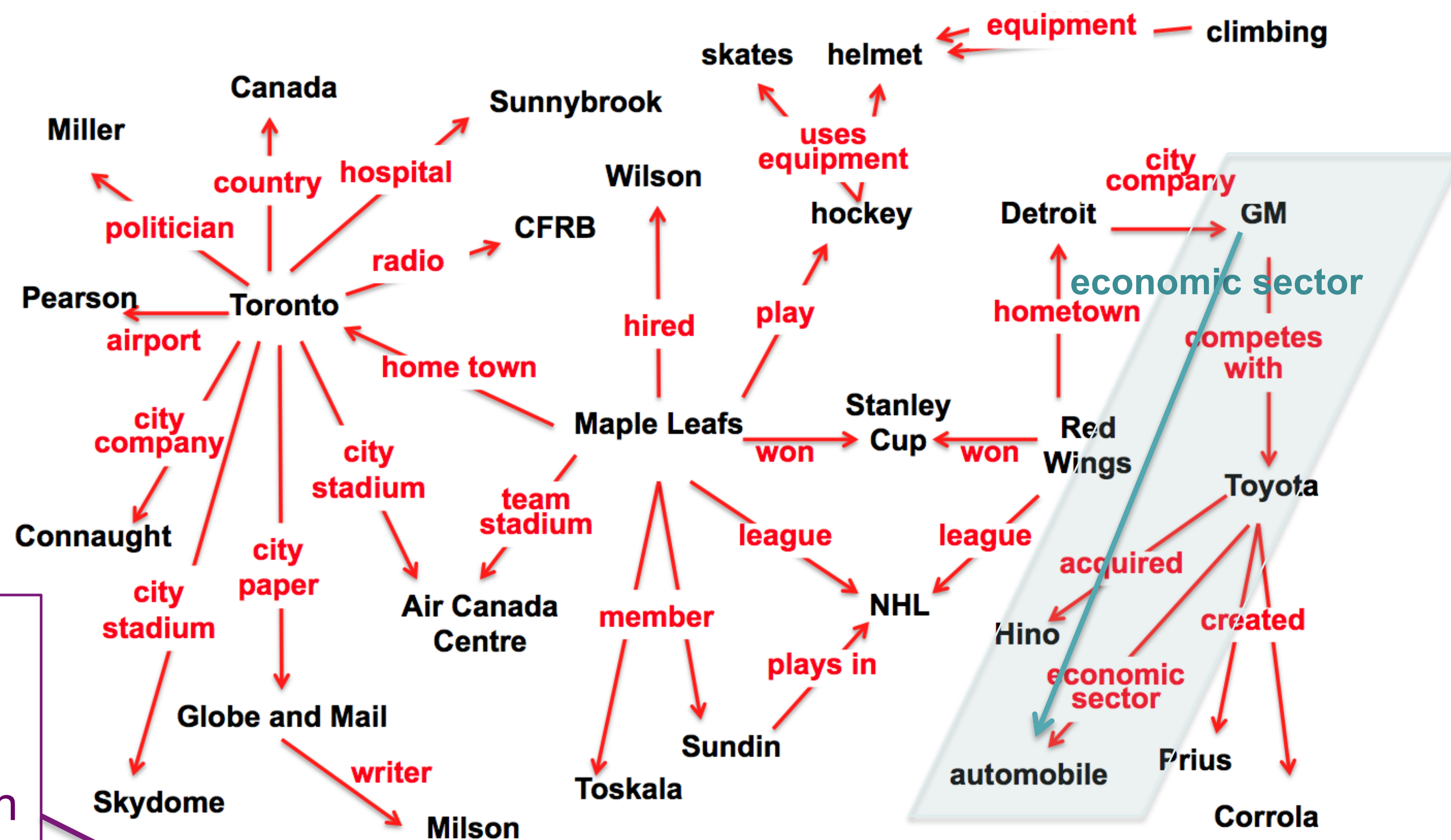


If: x1 — competes with (x1,x2) → x2 — economic sector (x2, x3) → x3

Then: economic sector (x1, x3)

Inference by Random Walks

PRA: [Lao, Mitchell, Cohen, *EMNLP* 2011]



PRA:

1. restrict precondition to a chain.

2. inference by random walks

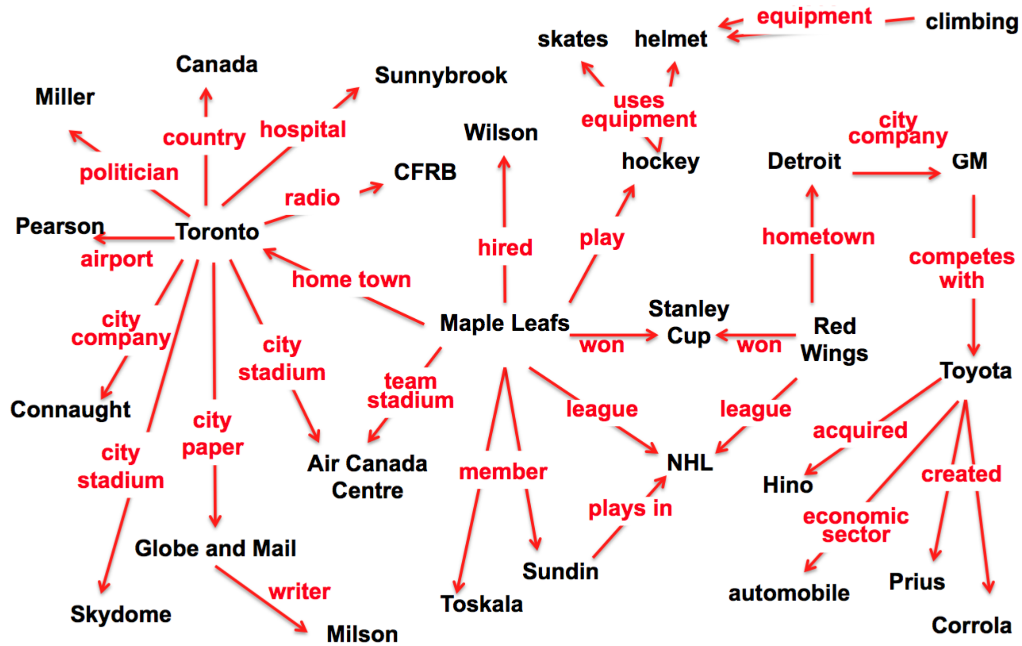
If: x1 — competes with (x1,x2) → x2 — economic sector (x2, x3) → x3

Then: economic sector (x1, x3)

Inference by KB Random Walks

[Lao, Mitchell, Cohen, *EMNLP* 2011]

KB:



Random walk path type: x $\xrightarrow{\text{competes with}}$? $\xrightarrow{\text{economic sector}}$ y

Pr(R(x,y)): logistic function for R(x,y)

where ith feature = probability of arriving at node y starting at node x, and taking a random walk along path of type i

CityLocatedInCountry(Pittsburgh) = ?

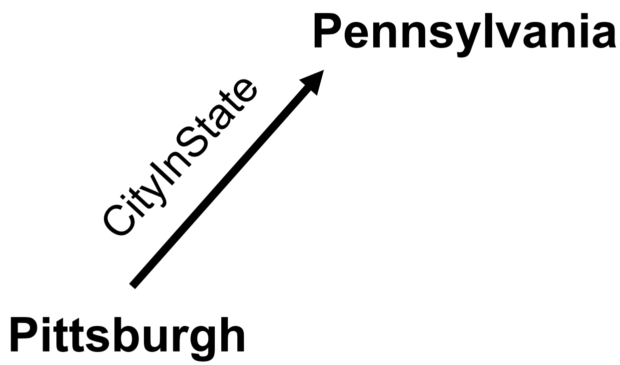
[Lao, Mitchell, Cohen, *EMNLP* 2011]

Pittsburgh

Feature = Typed Path	**Feature Value**	**Logistic Regresssion Weight**
CityInState, $CityInstate^{-1}$, CityLocatedInCountry		0.32

CityLocatedInCountry(Pittsburgh) = ?

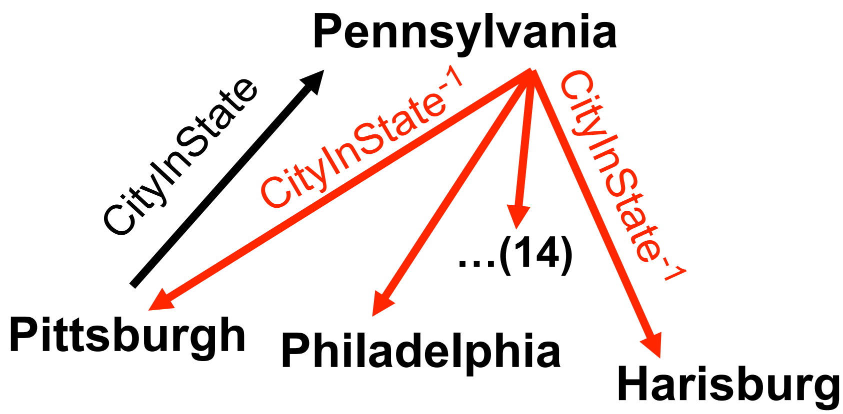
[Lao, Mitchell, Cohen, *EMNLP* 2011]



Feature = Typed Path	**Feature Value**	**Logistic Regresssion Weight**
CityInState, CityInstate^{-1}, CityLocatedInCountry		0.32

CityLocatedInCountry(Pittsburgh) = ?

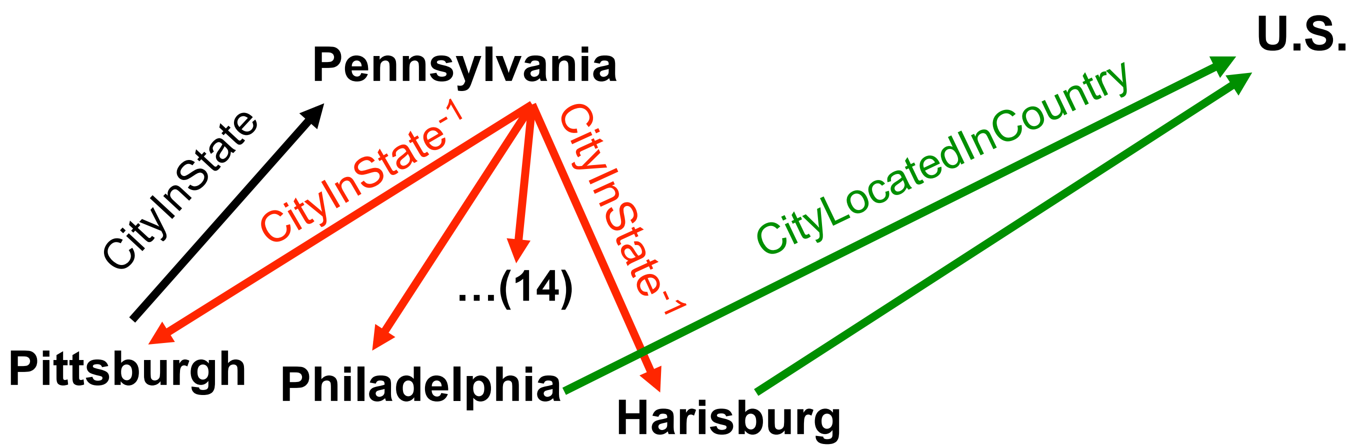
[Lao, Mitchell, Cohen, *EMNLP* 2011]



Feature = Typed Path	Feature Value	Logistic Regresssion Weight
CityInState, CityInstate^{-1}, CityLocatedInCountry		0.32

CityLocatedInCountry(Pittsburgh) = ?

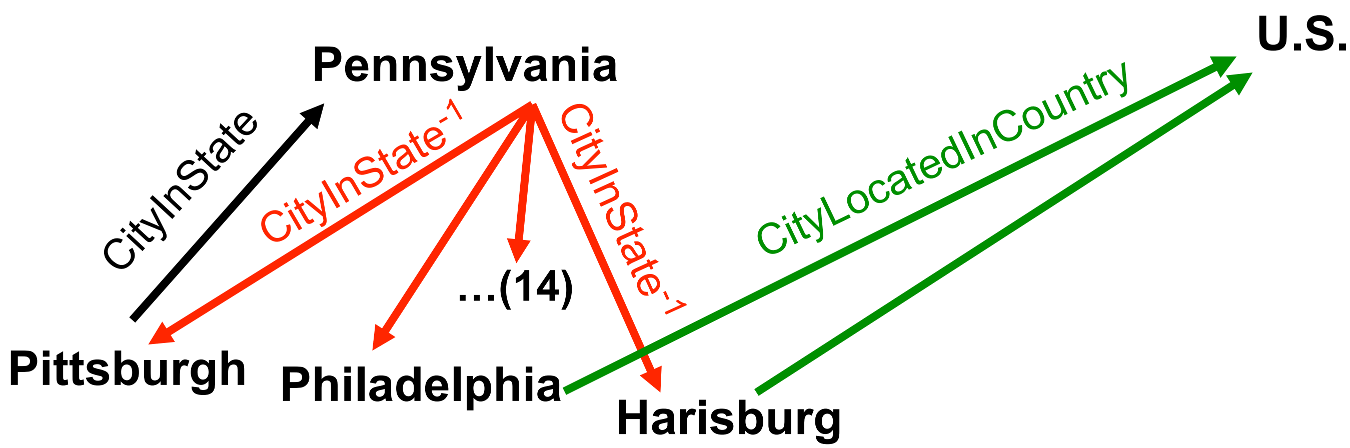
[Lao, Mitchell, Cohen, *EMNLP* 2011]



Feature = Typed Path	**Feature Value**	**Logistic Regresssion Weight**
CityInState, CityInstate^{-1}, CityLocatedInCountry		0.32

CityLocatedInCountry(Pittsburgh) = ?

[Lao, Mitchell, Cohen, *EMNLP* 2011]

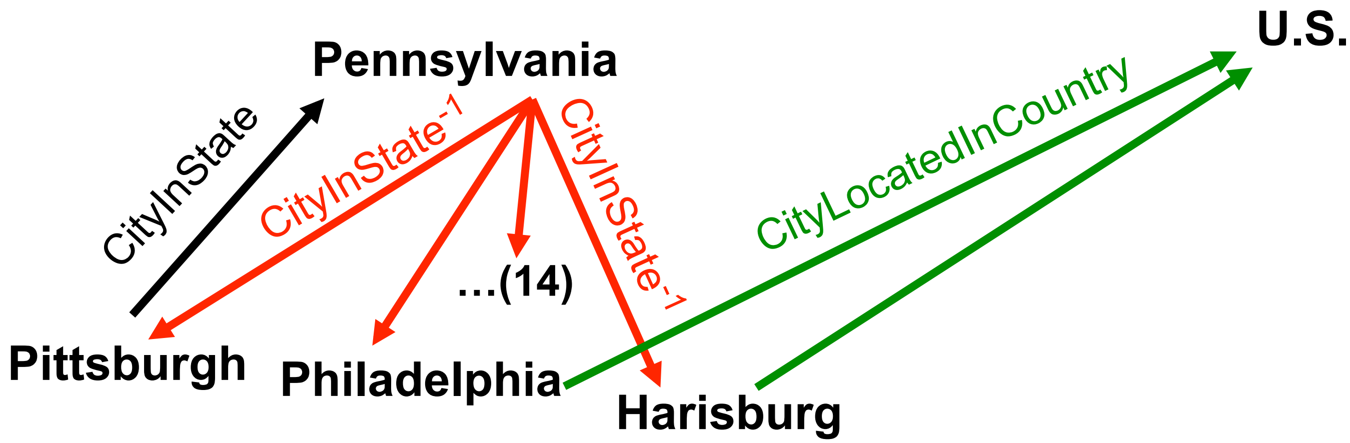


Pr(U.S. | Pittsburgh, TypedPath)

Feature = Typed Path	**Feature Value**	**Logistic Regresssion Weight**
CityInState, CityInstate^{-1}, CityLocatedInCountry	0.8	0.32

CityLocatedInCountry(Pittsburgh) = ?

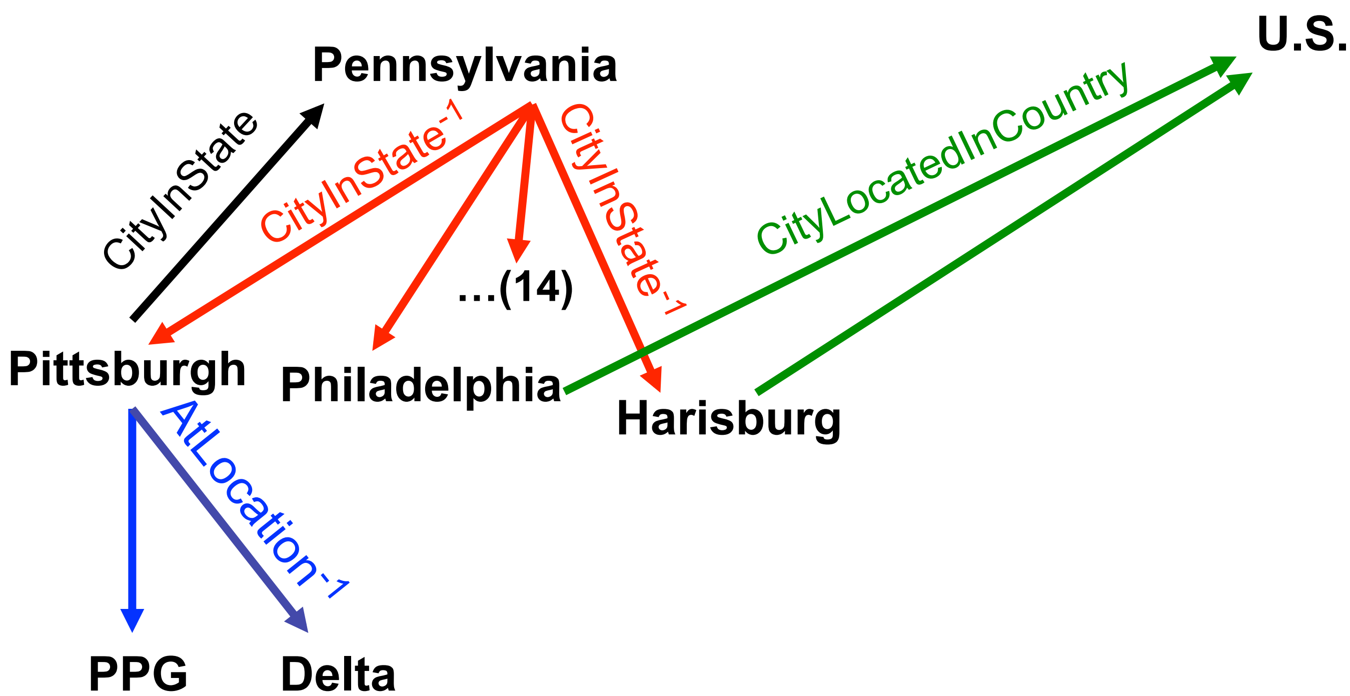
[Lao, Mitchell, Cohen, *EMNLP* 2011]



Feature = Typed Path	**Feature Value**	**Logistic Regresssion Weight**
CityInState, CityInstate^{-1}, CityLocatedInCountry	0.8	0.32
AtLocation^{-1}, AtLocation, CityLocatedInCountry		0.20

CityLocatedInCountry(Pittsburgh) = ?

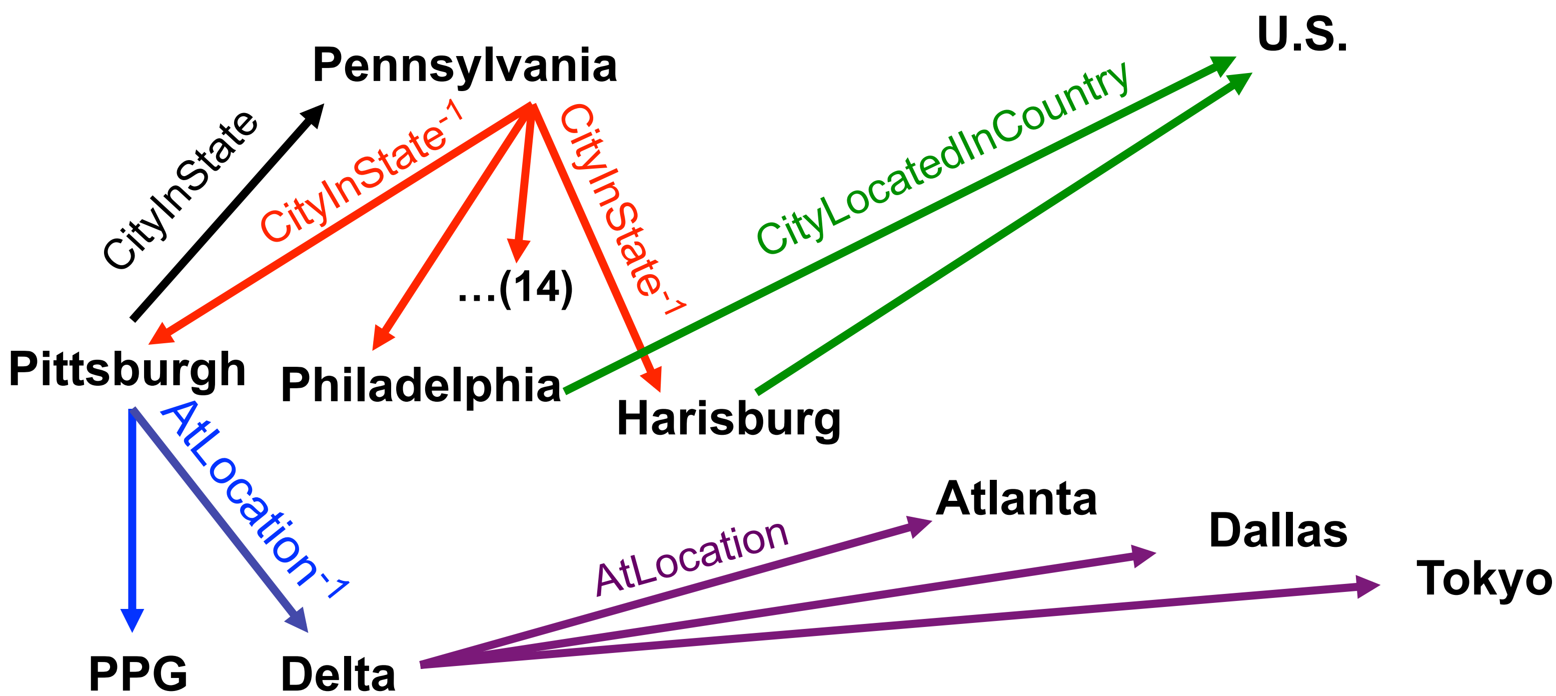
[Lao, Mitchell, Cohen, *EMNLP* 2011]



Feature = Typed Path	**Feature Value**	**Logistic Regresssion Weight**
CityInState, CityInstate^{-1}, CityLocatedInCountry	0.8	0.32
AtLocation^{-1}, AtLocation, CityLocatedInCountry		0.20

CityLocatedInCountry(Pittsburgh) = ?

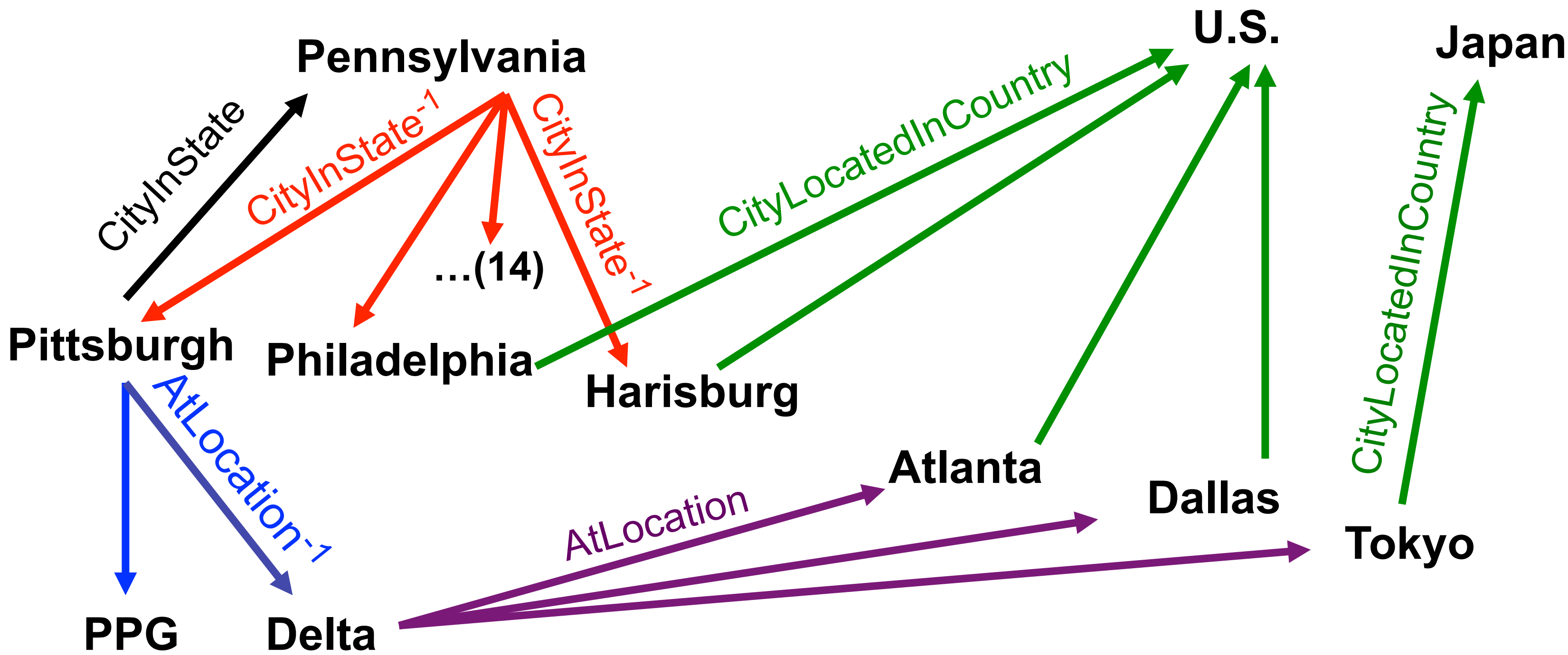
[Lao, Mitchell, Cohen, *EMNLP* 2011]



Feature = Typed Path	Feature Value	Logistic Regresssion Weight
CityInState, CityInstate^{-1}, CityLocatedInCountry	0.8	0.32
AtLocation^{-1}, AtLocation, CityLocatedInCountry		0.20

CityLocatedInCountry(Pittsburgh) = ?

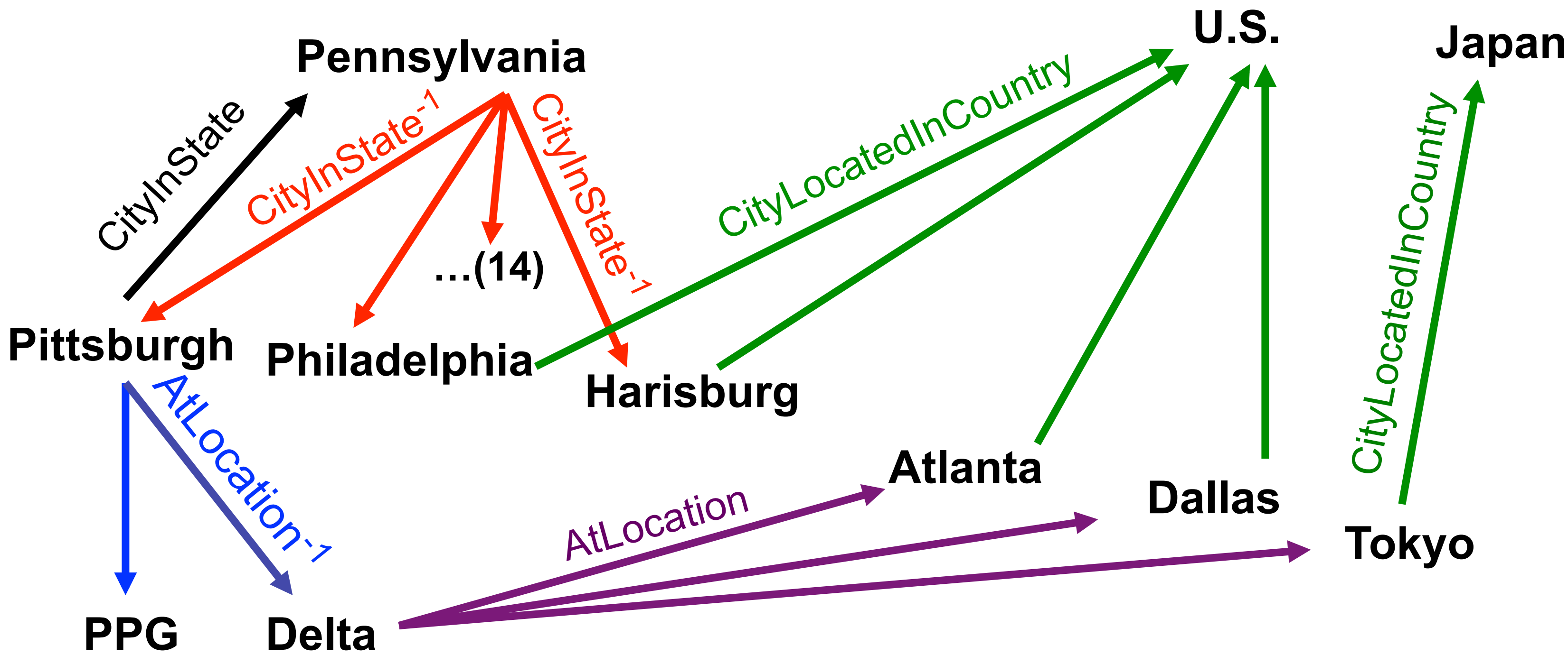
[Lao, Mitchell, Cohen, *EMNLP* 2011]



Feature = Typed Path	**Feature Value**	**Logistic Regresssion Weight**
CityInState, CityInstate^{-1}, CityLocatedInCountry	0.8	0.32
AtLocation^{-1}, AtLocation, CityLocatedInCountry	0.6	0.20

CityLocatedInCountry(Pittsburgh) = ?

[Lao, Mitchell, Cohen, *EMNLP* 2011]

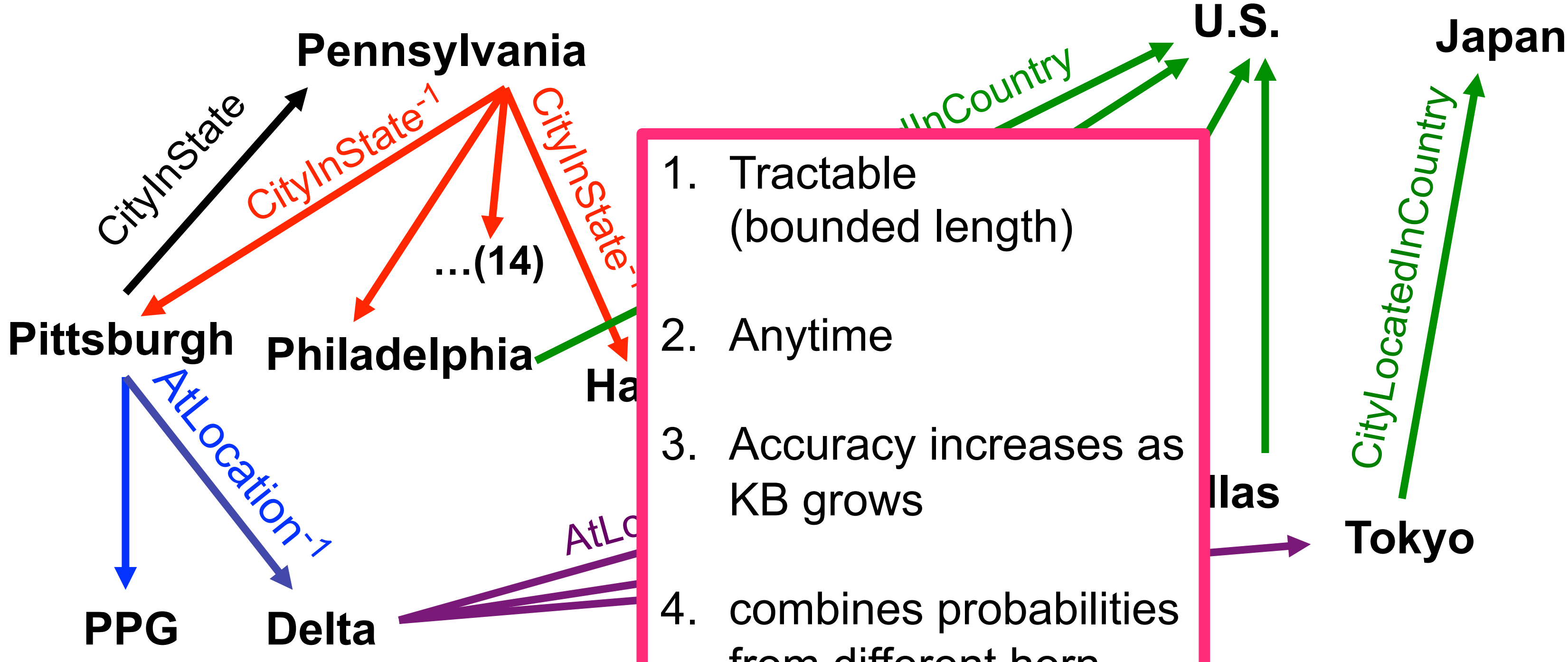


Feature = Typed Path	**Feature Value**	**Logistic Regresssion Weight**
CityInState, CityInstate^{-1}, CityLocatedInCountry	0.8	0.32
AtLocation^{-1}, AtLocation, CityLocatedInCountry	0.6	0.20
…	…	…

CityLocatedInCountry(Pittsburgh) = U.S.   p=0.58

CityLocatedInCountry(Pittsburgh) = ?

[Lao, Mitchell, Cohen, *EMNLP* 2011]



1. Tractable (bounded length)
2. Anytime
3. Accuracy increases as KB grows
4. combines probabilities from different horn clauses

Feature = Typed Path	**Feature Value**	**Logistic Regresssion Weight**
CityInState, CityInstate^{-1}, CityLocatedInCountry	0.8	0.32
AtLocation^{-1}, AtLocation, CityLocatedInCountry	0.6	0.20
…	…	…

CityLocatedInCountry(Pittsburgh) = U.S. p=0.58

Random walk inference: learned rules

CityLocatedInCountry(*city, country*):

8.04 cityliesonriver, cityliesonriver^{-1}, citylocatedincountry

5.42 hasofficeincity^{-1}, hasofficeincity, citylocatedincountry

4.98 cityalsoknownas, cityalsoknownas, citylocatedincountry

2.85 citycapitalofcountry,citylocatedincountry^{-1},citylocatedincountry

2.29 agentactsinlocation^{-1}, agentactsinlocation, citylocatedincountry

1.22 statehascapital^{-1}, statelocatedincountry

0.66 citycapitalofcountry

.

.

.

7 of the 2985 learned rules for CityLocatedInCountry

[Gardner et al, 2014]

Opportunity:

Can infer more if we start with more densely connected knowledge graph

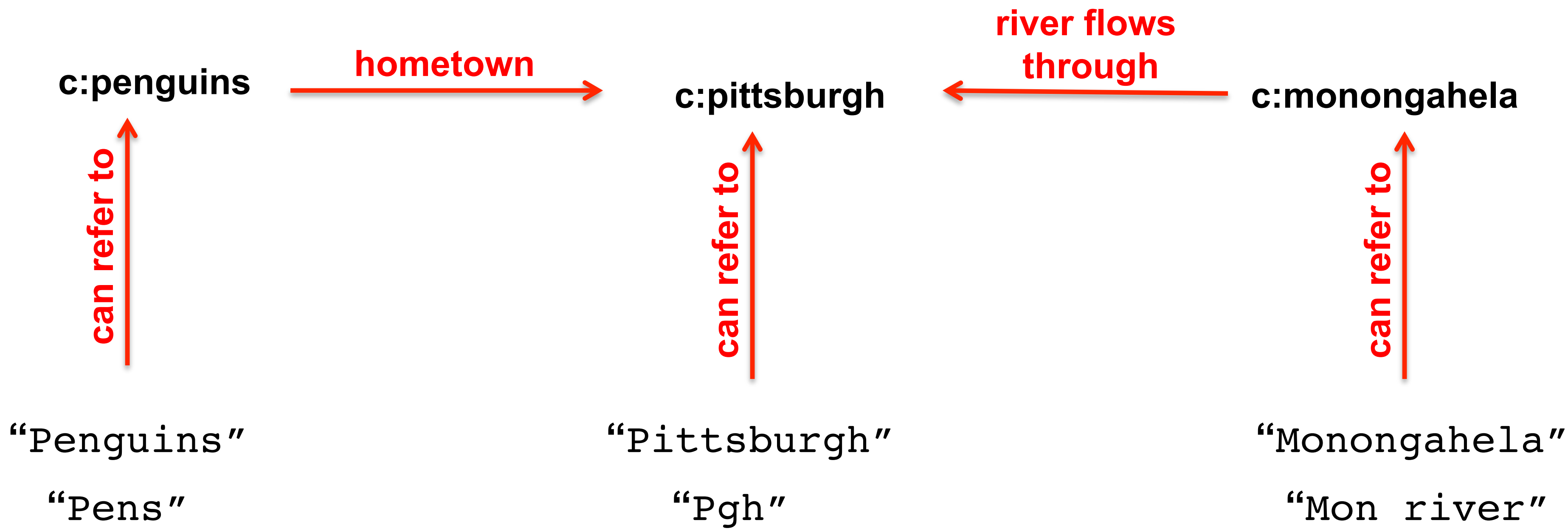
→ as NELL learns, it will become more dense

→ augment knowledge graph with a second graph of corpus statistics:

<subject, verb, object> triples

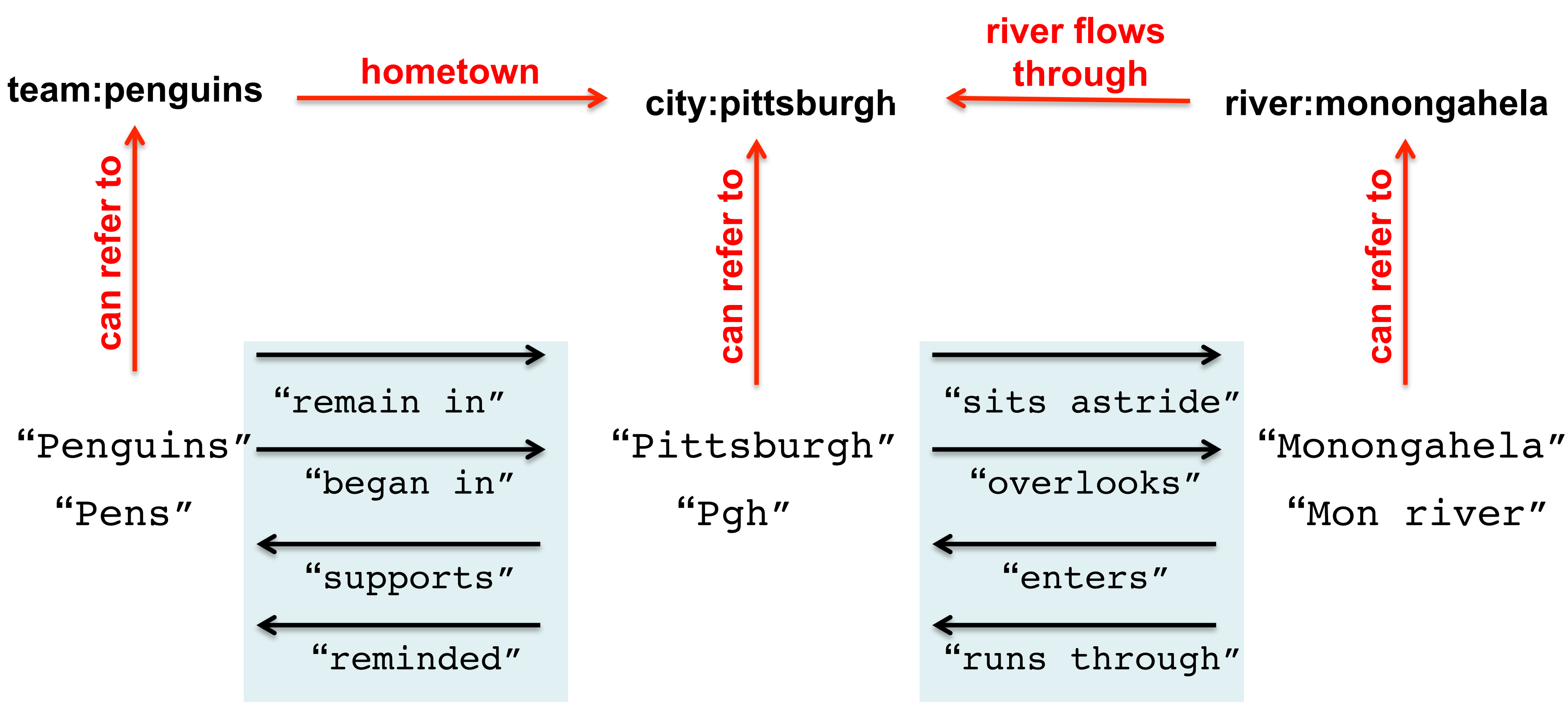
[Gardner et al, 2014]

NELL: concepts and "noun phrases"



[Gardner et al, 2014]

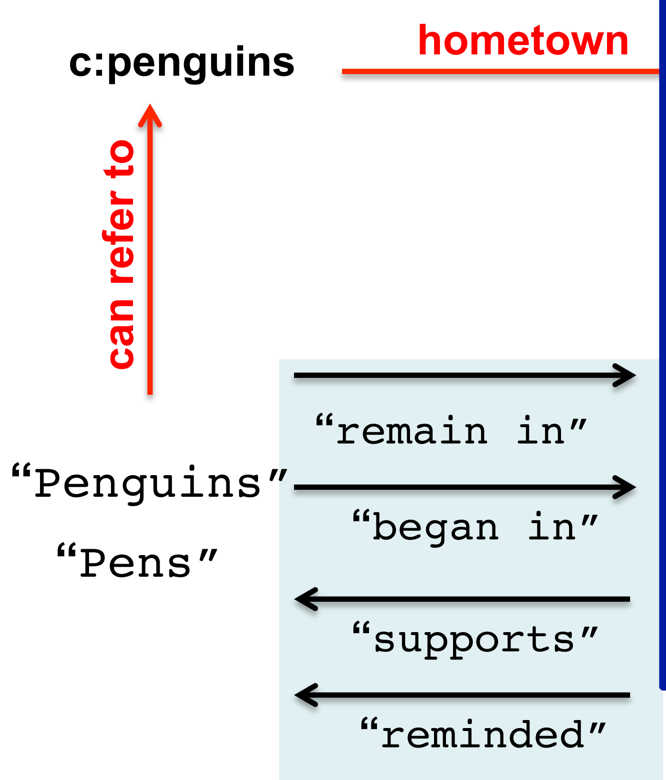
NELL: concepts and "noun phrases"



SVO triples from 500 M dependency parsed web pages (thank you Chris Re!)

[Gardner et al, 2014]

NELL: concepts and “noun phrases”



- Circumvents NELL’s fixed vocabulary of relations!
- Sadly, adding these does not help: too sparse
- But clustering verb phrases based on latent embedding (NNMF), produces significant improvement
    - {“lies on”, “runs through”, “flows through”, …}
- Precision/recall over 15 NELL relations:

KB only:	0.80 / 0.33
KB + SVO_{latent}:	0.87 / 0.42

[Gardner et al., 2014]

SVO triples from 500 M dependency parsed web pages (thank you Chris Re!)

Key Idea 3:

Automatically extend ontology

Ontology Extension (1)

[Mohamed et al., *EMNLP* 2011]

Goal:

- Add new relations to ontology

Approach:

- For each pair of categories C1, C2,
  - cluster pairs of known instances, in terms of text contexts that connect them

Example Discovered Relations

[Mohamed et al. *EMNLP* 2011]

Category Pair	Frequent Instance Pairs	Text Contexts	Suggested Name
MusicInstrument Musician	sitar, George Harrison tenor sax, Stan Getz trombone, Tommy Dorsey vibes, Lionel Hampton	ARG1 master ARG2 ARG1 virtuoso ARG2 ARG1 legend ARG2 ARG2 plays ARG1	Master
Disease Disease	pinched nerve, herniated disk tennis elbow, tendonitis blepharospasm, dystonia	ARG1 is due to ARG2 ARG1 is caused by ARG2	IsDueTo
CellType Chemical	epithelial cells, surfactant neurons, serotonin mast cells, histomine	ARG1 that release ARG2 ARG2 releasing ARG1	ThatRelease
Mammals Plant	koala bears, eucalyptus sheep, grasses goats, saplings	ARG1 eat ARG2 ARG2 eating ARG1	Eat
River City	Seine, Paris Nile, Cairo Tiber river, Rome	ARG1 in heart of ARG2 ARG1 which flows through ARG2	InHeartOf

NELL: sample of self-added relations

- athleteWonAward
- animalEatsFood
- languageTaughtInCity
- clothingMadeFromPlant
- beverageServedWithFood
- fishServedWithFood
- athleteBeatAthlete
- athleteInjuredBodyPart
- arthropodFeedsOnInsect
- animalEatsVegetable
- plantRepresentsEmotion
- foodDecreasesRiskOfDisease
- clothingGoesWithClothing
- bacteriaCausesPhysCondition
- buildingMadeOfMaterial
- emotionAssociatedWithDisease
- foodCanCauseDisease
- agriculturalProductAttractsInsect
- arteryArisesFromArtery
- countryHasSportsFans
- bakedGoodServedWithBeverage
- beverageContainsProtein
- animalCanDevelopDisease
- beverageMadeFromBeverage

Ontology Extension (2) [Burr Settles]

Goal:

- Add new subcategories

Approach:

- For each category C,
  - train NELL to read the relation $\text{SubsetOf}_C: C \rightarrow C$

*no new software here, just add this relation to ontology

NELL: subcategories discovered by reading

Animal:

- Pets
  - Hamsters, Ferrets, Birds, Dog, Cats, Rabbits, Snakes, Parrots, Kittens, …
- Predators
  - Bears, Foxes, Wolves, Coyotes, Snakes, Racoons, Eagles, Lions, Leopards, Hawks, Humans, …

Learned reading patterns for AnimalSubset(arg1,arg2)

"arg1 and other medium sized arg2" "arg1 and other jungle arg2" "arg1 and other magnificent arg2" "arg1 and other pesky arg2" "arg1 and other mammals and arg2" "arg1 and other Ice Age arg2" "arg1 or other biting arg2" "arg1 and other marsh arg2" "arg1 and other migrant arg2" "arg1 and other monogastric arg2" "arg1 and other mythical arg2" "arg1 and other nesting

NELL: subcategories discovered by reading

Animal:

- Pets
  - Hamsters, Ferrets, Birds, Dog, Cats, Rabbits, Snakes, Parrots, Kittens, …
- Predators
  - Bears, Foxes, Wolves, Coyotes, Snakes, Racoons, Eagles, Lions, Leopards, Hawks, Humans, …

Learned reading patterns:

"arg1 and other medium sized arg2" "arg1 and other jungle arg2" "arg1 and other magnificent arg2" "arg1 and other pesky arg2" "arg1 and other mammals and arg2" "arg1 and other Ice Age arg2" "arg1 or other biting arg2" "arg1 and other marsh arg2" "arg1 and other migrant arg2" "arg1 and other monogastric arg2" "arg1 and other mythical arg2" "arg1 and other nesting

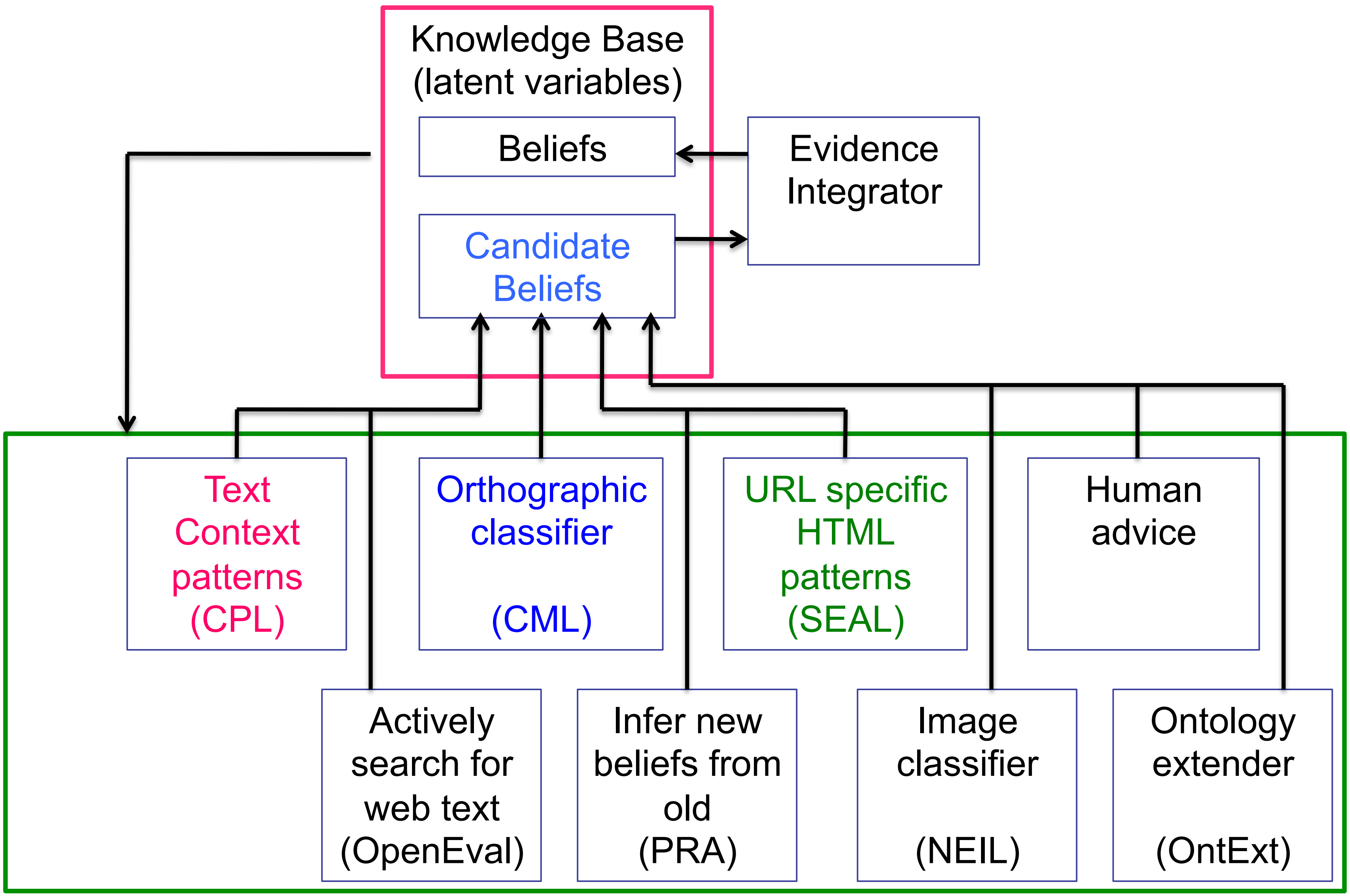
Chemical:

- Fossil fuels
  - Carbon, Natural gas, Coal, Diesel, Monoxide, Gases, …
- Gases
  - Helium, Carbon dioxide, Methane, Oxygen, Propane, Ozone, Radon…

Learned reading patterns:

"arg1 and other hydrocarbon arg2" "arg1 and other aqueous arg2" "arg1 and other hazardous air arg2" "arg1 and oxygen are arg2" "arg1 and such synthetic arg2" "arg1 as a lifting arg2" "arg1 as a tracer arg2" "arg1 as the carrier arg2" "arg1 as the inert arg2" "arg1 as the primary cleaning arg2" "arg1 and other noxious arg2" "arg1 and other trace arg2" "arg1 as the reagent arg2" "arg1 as the tracer

NELL Architecture



Key Idea 4: Cumulative, Staged Learning

Learning X improves ability to learn Y

1. Classify noun phrases (NP's) by category
2. Classify NP pairs by relation
3. Discover rules to predict new relation instances
4. Learn which NP's (co)refer to which latent concepts
5. Discover new relations to extend ontology
6. Learn to infer relation instances via targeted random walks
7. Vision: connect NELL and NEIL

NELL is here

8. Learn to microread single sentences
9. Learn to assign temporal scope to beliefs
10. Goal-driven reading: predict, then read to corroborate/correct
11. Make NELL a conversational agent on Twitter
12. Add a robot body to NELL

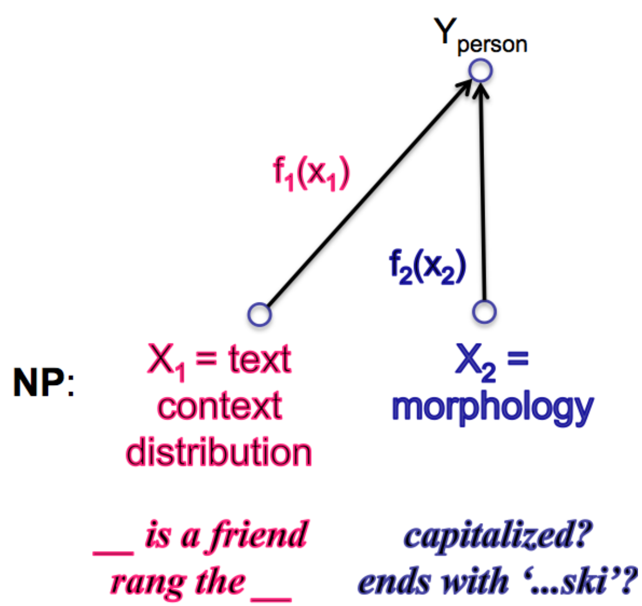
Consistency
Correctness
Self reflection

The core problem:

- Agents can measure internal *consistency*, but not *correctness*

Challenge:

- Under what conditions does *consistency* → *correctness*?



The core problem:

- Agents can measure internal *consistency*, but not *correctness*

Challenge:

- Under what conditions does *consistency* → *correctness*?
- Can an autonomous agent determine its accuracy from observed consistency?

[Platanios, Blum, Mitchell, *UAI 2014*]

Problem setting:

- have N different estimates $f_1, \ldots f_N$ of target function f^*

  $f_i : X \rightarrow Y; \quad Y \in \{0,1\}$

- *agreement* between f_i, f_j : $a_{ij} \equiv P_x(f_i(x) = f_j(x))$

[Platanios, Blum, Mitchell, *UAI 2014*]

Problem setting:

- have N different estimates $f_1, \ldots f_N$ of target function f^*
  $f_i : X \rightarrow Y; \quad Y \in \{0, 1\}$
- *agreement* between f_i, f_j : $a_{ij} \equiv P_x(f_i(x) = f_j(x))$

Key insight: errors and agreement rates are related

$$a_{ij} = \text{Pr[neither makes error] + Pr[both make error]}$$

$$a_{ij} = 1 - e_i - e_j + 2e_{ij}$$

prob. f_i and f_i agree

prob. f_i error

prob. f_j error

prob. f_i and f_j *both* make error

Estimating Error from Unlabeled Data

1. IF f_1, f_2, f_3 make indep. errors, and accuracies > 0.5

   THEN $$a_{ij} = 1 - e_i - e_j + 2e_{ij}$$

   → $$a_{ij} = 1 - e_i - e_j + 2e_i e_j$$

Measure errors from unlabeled data:

- use unlabeled data to estimate a_{12}, a_{13}, a_{23}
- solve three equations for three unknowns e_1, e_2, e_3

Estimating Error from Unlabeled Data

1. IF f_1, f_2, f_3 make indep. errors, accuracies > 0.5

   THEN $a_{ij} = 1 - e_i - e_j + 2e_{ij}$

   → $a_{ij} = 1 - e_i - e_j + 2e_i e_j$

2. but if errors not independent

Estimating Error from Unlabeled Data

1. IF f_1, f_2, f_3 make indep. errors, accuracies > 0.5

   THEN $a_{ij} = 1 - e_i - e_j + 2e_{ij}$

   → $a_{ij} = 1 - e_i - e_j + 2e_i e_j$

2. but if errors not independent

$$\min \quad (e_{ij} - e_i e_j)^2$$

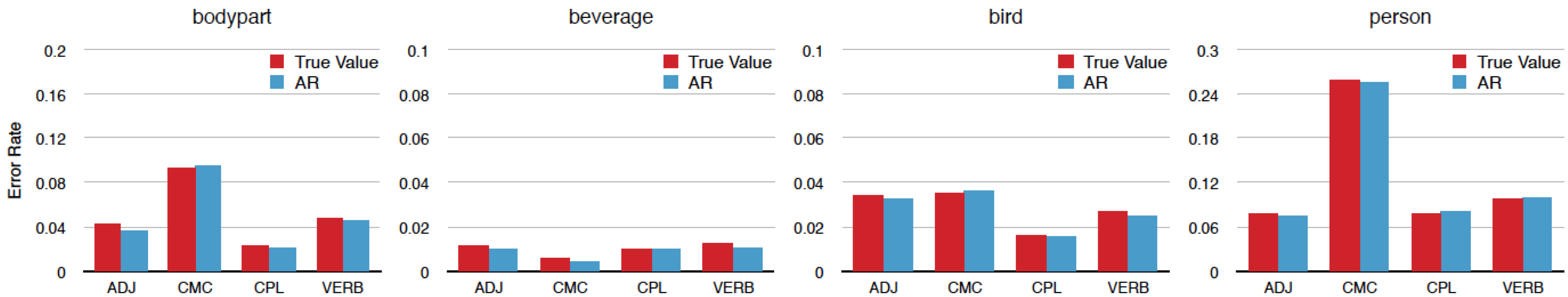
such that

$$(\forall i, j)\ a_{ij} = 1 - e_i - e_j + 2e_{ij}$$

True error (red), estimated error (blue)

[Platanios, Blum, Mitchell, *UAI 2014*]

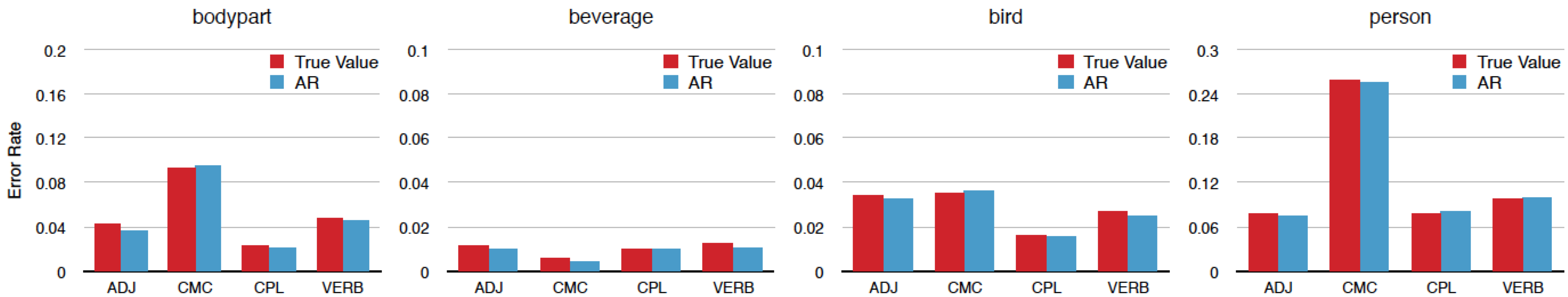
NELL classifiers:



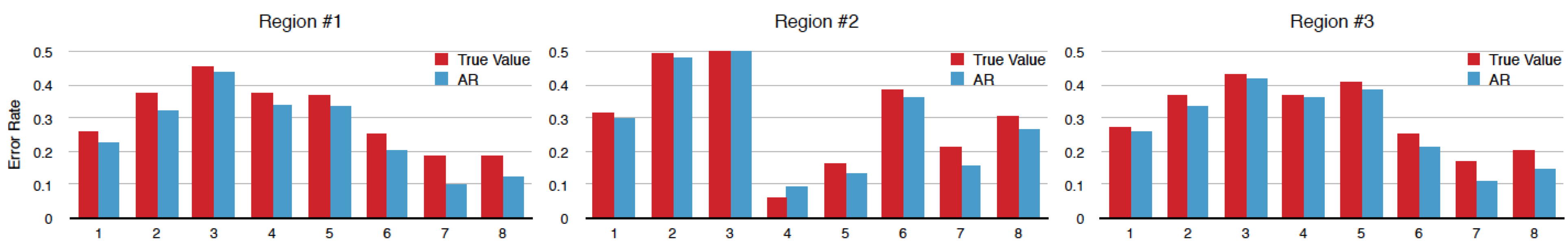
True error (red), estimated error (blue)

[Platanios, Blum, Mitchell, *UAI 2014*]

NELL classifiers:



Brain image fMRI classifiers:



Summary

1. Use *coupled* training for semi-supervised learning
2. Datamine the KB to learn probabilistic inference rules
3. Automatically extend ontology
4. Use staged learning curriculum

New directions:

- Self-reflection, self-estimates of accuracy (A. Platanios)
- Incorporate vision with NEIL (Abhinav Gupta)
- Microreading (Jayant Krishnamurthy, Ndapa Nakashole)
- Aggressive ontology expansion (Derry Wijaya)
- Portuguese NELL (Estevam Hrushka)
- never-ending learning phones? robots? traffic lights?

thank you

and thanks to:
Darpa, Google, NSF, Yahoo!, Microsoft, Fulbright, Intel

follow NELL on Twitter: @CMUNELL
browse/download NELL's KB at http://rtw.ml.cmu.edu