Statistical Relational Learning and Knowledge Graph Reasoning

CSCI 699

Jay Pujara
Reminder: Basic problems

• **Who** are the entities (nodes) in the graph?

• **What** are their attributes and types (labels)?

• **How** are they related (edges)?
Motivating Problem: New Opportunities

- Internet
  - Massive source of publicly available information
- Extraction
  - Cutting-edge IE methods
- Knowledge Graph (KG)
  - Structured representation of entities, their labels and the relationships between them
Motivating Problem: Real Challenges

Internet

Noisy!

Extraction

Difficult!

Knowledge Graph

Contains many errors and inconsistencies
Graph Construction Issues

Extracted knowledge is:

- ambiguous:
  - Ex: Beetles, beetles, Beatles
  - Ex: citizenOf, livedIn, bornIn
Graph Construction Issues

Extracted knowledge is:

- ambiguous

- incomplete
  - Ex: missing relationships
  - Ex: missing labels
  - Ex: missing entities
Graph Construction Issues

Extracted knowledge is:

- ambiguous
- incomplete
- inconsistent
  - Ex: Cynthia Lennon, Yoko Ono
  - Ex: exclusive labels (alive, dead)
  - Ex: domain-range constraints
Graph Construction Issues

Extracted knowledge is:

- ambiguous
- incomplete
- inconsistent
NELL: The Never-Ending Language Learner

- Large-scale IE project (Carlson et al., 2010)
- Lifelong learning: aims to “read the web”
- Ontology of known labels and relations
- Knowledge base contains millions of facts
Examples of NELL errors
Kyrgyzstan has many variants:
- Kyrgyzstan
- Kyrgistan
- Kyrghyzstan
- Kyrgyzstan
- Kyrgyz Republic

Saudi Cultural Days in the Kyrgyz Republic has concluded its activities in the capital Bishkek in the weekend in a special ceremony held on this occasion. The event was attended by Deputy Minister of Culture and Tourism of the Kyrgyz Republic Koulev Mirza; Kyrgyzstan’s Ambassador to Saudi Arabia Jusupbek Sharipov; the Saudi Embassy Acting Chargé d’affaires to Kyrgyzstan, Mari bin Barakah Al-Derbas and members of the embassy staff, in the presence of a heavy turnout of Kyrgyz citizens.

The Days of Culture of Saudi Arabia in Kyrgyzstan will be held from 6 to 9 May.

Refugees are often from areas where conflict is historically embedded and marked in ideology and injustice. The Tsarnaev family emigrated from the Chechen diaspora in Kyrgzstan, a region Stalin deported the Chechens to in 1943. After the fall of the Berlin Wall in 1991, Chechens engaged in a battle for independence from Russia that led to the Tsarnaevs' petition for refugee status in the early
Missing and spurious labels

Anssi Kullberg has sent along some great trip reports to unusual places, including Kyrgyzstan, Pakistan, Egypt/Jordan, and Afghanistan. I had to create a whole new country page for Afghanistan to hold that last one! Thanks so much, Anssi!

Erik Kleyheeg has just returned from Lesvos with some new bird images. Included here are: Common Scops-Owl, Wood Warbler, Spanish Sparrow, Red-throated Pipit, Eurasian Chiff-chaff, and Cretzschmar's Bunting.

Kyrgyzstan (/kərˈɡɪstɑːn/ kər-gi-stahn;[5] Kyrgyz: Кыргызстан (IPA: [qyrɡyʃˈstan]); Russian: Киргизия), officially the Kyrgyz Republic (Kyrgyz: Кыргыз Республикасы; Russian: Кыргызская Республика), is a country located in Central Asia.[6] Landlocked and mountainous, Kyrgyzstan is bordered by Kazakhstan to the north, Uzbekistan to the west, Tajikistan to the southwest and China to the east. Its capital and largest city is Bishkek.
Missing and spurious relations

Guidance

Kazakhstan / Kyrgyzstan – Consular Fees

Kyrgyzstan’s location is ambiguous – Kazakhstan, Russia and US are included in possible locations

Kyrgyzstan U.S. Air Base Future Unclear

A Central Asian country of incredible natural beauty and proud nomadic traditions, most of Kyrgyzstan was formally annexed to Russia in 1876. The Kyrgyz staged a major revolt against the Tsarist Empire in 1916 in which almost one-sixth of the Kyrgyz population was killed. Kyrgyzstan became a Soviet republic in 1936 and
Violations of ontological knowledge

• Equivalence of co-referent entities (sameAs)
  • SameEntity(Kyrgyzstan, Kyrgyz Republic)

• Mutual exclusion (disjointWith) of labels
  • MUT(bird, country)

• Selectional preferences (domain/range) of relations
  • RNG(countryLocation, continent)

Enforcing these constraints require jointly considering multiple extractions
Graph Construction approach

- Graph construction cleans and completes extraction graph

- Incorporate ontological constraints and relational patterns

- Discover statistical relationships within knowledge graph
Graph Construction
Probabilistic Models

TOPICS:

OVERVIEW

GRAPHICAL MODELS

RANDOM WALK METHODS
Graph Construction
Probabilistic Models

TOPICS:

OVERVIEW

GRAPHICAL MODELS

RANDOM WALK METHODS
Voter Party Classification
Voter Party Classification

Multiple Sources of Information
Voter Party Classification

Multiple Sources of Information

- Statuses & Tweets
- Donations
Voter Party Classification

Multiple Sources of Information

- Statuses & Tweets
- Donations
- Friends & Followers
Voter Party Classification

Multiple Sources of Information

- Statuses & Tweets
- Donations
- Friends & Followers
- Family
Voter Party Classification

@mrtimlong

Forget Trump & his 100's of accusers — this new Hillary REVELATION will turn the race UPSIDE DOWN!!!
Voter Party Classification

sliderwave @sliderwave · Oct 15
@mrtimlong @ABC7NY OMG! Someone in the campaign was ready and prepared because it's the most important job on Earth? Why would they do that?

Emily Fun Buns @MsEffieLou · Oct 15
@mrtimlong almost as bad as her sympathizing with Bernie supporters and not calling them Basement Dwellers

(((Voter))) Eitan @AnotherEitan · Oct 15
@mrtimlong @ABC7NY Holy shit. I'm starting to think that she Wants to be President.
Voter Party Classification

Tim Long
@mrtimlong
Simpsons Writer/ Man of Peace
Location: Los Angeles
Joined: December 2010

Carly Fiorina for Vice President.com
Voter Party Classification

Tim Long
@mrtimlong
Simpsons Writer/ Man of Peace
Location: Los Angeles
Joined: December 2010

Donate to the Planned Parenthood Action Fund.

Tell your friends and family!
Voter Party Classification

Multiple Sources of Information

- Statues & Tweets
- Donations
- Friends & Followers
- Family
Standard Classification

Forget Trump & his 100’s of accusers — this new Hillary REVELATION will turn the race UPSIDE DOWN!!!

CarlyFiorinaforVicePresident.com

Bag-of-words features
Standard Classification

Forget Trump & his 100’s of accusers — this new Hillary REVELATION will turn the race UPSIDE DOWN!!!

Eyewitness News @ABC7NY
Clinton Team Ran Highly Scripted Campaign, WikiLeaks Emails Indicate 7ny.tv/2deCcrL

CarlyFiorinaforVicePresident.com

Bag-of-words features

Pr(Y)
Standard Classification

CarlyFiorinaforVicePresident.com

Bag-of-words features
Voter Party Classification

Multiple Sources of Information

Donations

Status Updates

Friends

Family
Collective Classification

Tim Long
@mrtimlong
Simpsons Writer/ Man of Peace
Location: Los Angeles
Joined: December 2010

Terri Lee Dee
@terrileedee

Follows
Collective Classification
Collective Classification

Tim Long
@mrtimlong
Simpsons Writer/ Man of Peace
📍 Los Angeles
📅 Joined December 2010

Terri Lee Dee
@terrileeddee

Pr(Y): [Bar Chart]

Follows

WOMEN CAN
STOP TRUMP
Collective Classification

My label is likely to match that of my follower
Collective Classification

Follows(U₁, U₂) & Votes(U₁, P) → Votes(U₂, P)
Collective Classification

spouse

follower
Collective Classification

spouse  follower
Collective Classification

Spouse(U1, U2) & Votes(U1, P) $\rightarrow$ Votes(U2, P)

Follows(U1, U2) & Votes(U1, P) $\rightarrow$ Votes(U2, P)
Collective Classification

Wait don't forget that Hillary caught a bug & stayed home for THREE DAYS

Don't blame me, I voted for Kodos.

They both have problems.
Collective Classification
Collective Classification

spouse

follower

follower

follower
Collective Classification

5.0: Spouse(U1, U2) \& Votes(U1, P) \rightarrow Votes(U2, P)

2.0: Follows(U1, U2) \& Votes(U1, P) \rightarrow Votes(U2, P)
Collective Classification
Collective Classification
Collective Classification with PSL

/* Local rules */
5.0: Donates(A, P) -> Votes(A, P)
0.3: Mentions(A, "Affordable Health") -> Votes(A, "Democrat")
0.3: Mentions(A, "Tax Cuts") -> Votes(A, "Republican")

/* Relational rules */
1.0: Votes(A,P) & Spouse(B,A) -> Votes(B,P)
0.3: Votes(A,P) & Friend(B,A) -> Votes(B,P)
0.1: Votes(A,P) & Colleague(B,A) -> Votes(B,P)

/* Range constraint */
Votes(A, "Republican") + Votes(A, "Democrat") = 1.0.
Beyond Pure Reasoning

- Classical AI approach to knowledge: reasoning

\[ \text{Lbl}(\text{Socrates, Man}) \land \text{Sub}(\text{Man, Mortal}) \Rightarrow \text{Lbl}(\text{Socrates, Mortal}) \]
Beyond Pure Reasoning

• Classical AI approach to knowledge: reasoning

Lbl(Socrates, Man) & Sub(Man, Mortal) -> Lbl(Socrates, Mortal)

• Reasoning difficult when extracted knowledge has errors
Beyond Pure Reasoning

- Classical AI approach to knowledge: reasoning
  \[\text{Lbl}(\text{Socrates, Man}) \land \text{Sub}(\text{Man, Mortal}) \to \text{Lbl}(\text{Socrates, Mortal})\]
- Reasoning difficult when extracted knowledge has errors
- Solution: probabilistic models
  \[P(\text{Lbl}(\text{Socrates, Mortal})|\text{Lbl}(\text{Socrates, Man})=0.9)\]
Logic Refresher: Satisfaction

/* Model Snippet */
Mentions(A, “Affordable Health”) -> Votes(A, “Democrat”)

<table>
<thead>
<tr>
<th>Affordable Health</th>
<th>Democrat</th>
<th>Logical Satisfaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRUE</td>
<td>TRUE</td>
<td>😊</td>
</tr>
<tr>
<td>TRUE</td>
<td>FALSE</td>
<td>😞</td>
</tr>
<tr>
<td>FALSE</td>
<td>TRUE</td>
<td>😊</td>
</tr>
<tr>
<td>FALSE</td>
<td>FALSE</td>
<td>😊</td>
</tr>
</tbody>
</table>
Logic and Noisy Data

/* Model Snippet */

[1] Mentions(A, "Affordable Health") -> Votes(A, "Democrat")

<table>
<thead>
<tr>
<th>Affordable Health</th>
<th>Tax Cuts</th>
<th>Democrat</th>
<th>[1] Logical Satisfaction</th>
<th>[2] Logical Satisfaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRUE</td>
<td>TRUE</td>
<td>TRUE</td>
<td><img src="emoji" alt="green" /></td>
<td><img src="emoji" alt="red" /></td>
</tr>
<tr>
<td>TRUE</td>
<td>TRUE</td>
<td>FALSE</td>
<td><img src="emoji" alt="red" /></td>
<td><img src="emoji" alt="green" /></td>
</tr>
</tbody>
</table>
Logic and Noisy Data

/* Model Snippet */

[1] \( \text{Mentions}(A, \text{“Affordable Health”}) \rightarrow \text{Votes}(A, \text{“Democrat”}) \)

[2] \( \text{Mentions}(A, \text{“Tax Cuts”}) \rightarrow \neg \text{Votes}(A, \text{“Democrat”}) \)

<table>
<thead>
<tr>
<th>Affordable Health</th>
<th>Tax Cuts</th>
<th>Democrat</th>
<th>[1] Logical Satisfaction</th>
<th>[2] Logical Satisfaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRUE</td>
<td>TRUE</td>
<td>TRUE</td>
<td>☺</td>
<td>☹</td>
</tr>
<tr>
<td>TRUE</td>
<td>TRUE</td>
<td>FALSE</td>
<td>☹</td>
<td>☻</td>
</tr>
</tbody>
</table>

In logic, much as in politics, it is hard to satisfy everyone.
Soft Logic to the Rescue!

/* Model Snippet */

<table>
<thead>
<tr>
<th>Affordable Health</th>
<th>Tax Cuts</th>
<th>Democrat</th>
<th>[1] Logical Satisfaction</th>
<th>[2] Logical Satisfaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRUE</td>
<td>TRUE</td>
<td>0.5</td>
<td>😞</td>
<td>😞</td>
</tr>
<tr>
<td>TRUE</td>
<td>TRUE</td>
<td>0.5</td>
<td>😞</td>
<td>😞</td>
</tr>
</tbody>
</table>
What does 0.5 MEAN?
What does 0.5 mean?

• Rounding probability:
  • Flip a coin with bias 0.5
  • Heads = TRUE
  • Tails = FALSE

• Using this method is a ¾ optimal solution to the NP hard weighted MAX SAT problem
  [Goemans&Williams, 94]
What does 😐 MEAN?
What does 🤔 mean?

\[ P \rightarrow Q \]

• /* Soft Logic Penalty */

• if \( P < Q \)
  
  return 😊

• else:
  • return \( P - Q \)
Closed Form

$\max(0, P - Q)$
$P \rightarrow Q$

$\max(0, P-Q)$

🤔: Closed Form
What does 😐 mean?

/* Model Snippet */

/* Soft Logic Penalty */

    return 0
else:
/* Model Snippet */


<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0.2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[
!Q = 1 - Q \\
P \rightarrow Q = \max(0, P - Q)
\]
/* Model Snippet */

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0.7</td>
<td>0.3</td>
<td>0.7</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0.2</td>
<td>0.8</td>
<td>0.2</td>
</tr>
</tbody>
</table>

!Q = 1-Q
P -> Q = max(0, P-Q)
Computing 😞 with soft evidence

/* Model Snippet */


<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4</td>
<td>0.1</td>
<td>0.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td>0.1</td>
<td>0.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td>0.1</td>
<td>0.9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

!Q = 1-Q

P -> Q = max(0, P-Q)
Computing 😞 with soft evidence

/* Model Snippet */

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4</td>
<td>0.1</td>
<td>0.65</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.4</td>
<td>0.1</td>
<td>0.2</td>
<td>0.2</td>
<td>0</td>
</tr>
<tr>
<td>0.4</td>
<td>0.1</td>
<td>0.9</td>
<td>0.5</td>
<td>0</td>
</tr>
</tbody>
</table>

!Q = 1-Q
P -> Q = max(0, P-Q)
Computing 😞 for arbitrary formulas

\[！Q = 1 - Q\]
\[P \rightarrow Q = \max(0, P - Q)\]
\[P \& Q = \max(0, P + Q - 1)\]
\[P | Q = \min(1, P + Q)\]
Underlying Probability Distribution

\[
p(Y|X) = \frac{1}{Z(w, X)} \exp \left[ - \sum_{j=1}^{m} w_j \left[ \max \{ \ell_j(Y, X), 0 \} \right]^{1,2} \right]
\]

- Joint probability over soft-truth assignments
- Sum over rule penalties
Optimizing PSL models

• PSL finds optimal assignment for all unknowns

• Optimal = minimizes the soft-logic penalty

• **Fast**, joint convex optimization using ADMM

• Supports learning rule weights and latent variables
Graph Construction
Probabilistic Models

TOPICS:

OVERVIEW

GRAPHICAL MODELS

RANDOM WALK METHODS
Graphical Models: Overview

- Define **joint probability distribution** on knowledge graphs

- Each candidate fact in the knowledge graph is a **variable**

- Statistical signals, ontological knowledge and rules parameterize the **dependencies** between variables

- Find most likely knowledge graph by **optimization/sampling**
Motivating Problem (revised)

Internet → (noisy) Extraction Graph → Large-scale IE → Knowledge Graph

= Joint Reasoning
Knowledge Graph Identification

Problem:

Extraction Graph

Knowledge Graph Identification

Solution: Knowledge Graph Identification (KGI)

Performs graph identification:
- entity resolution
- collective classification
- link prediction

Enforces ontological constraints

Incorporates multiple uncertain sources
Define a graphical model to perform all three of these tasks simultaneously!

• **Who** are the entities (nodes) in the graph?

• **What** are their attributes and types (labels)?

• **How** are they related (edges)?
Knowledge Graph Identification

P(Who, What, How | Extractions)
Probabilistic Models

- Use dependencies between facts in KG

- Probability defined *jointly* over facts
What determines probability?

• Statistical signals from text extractors and classifiers
What determines probability?

- **Statistical signals from text extractors and classifiers**
  - \( P(R(\text{John, Spouse, Yoko})) = 0.75; \ P(R(\text{John, Spouse, Cynthia})) = 0.25 \)
  - \( \text{LevenshteinSimilarity} (\text{Beatles, Beetles}) = 0.9 \)
What determines probability?

- Statistical signals from text extractors and classifiers
- Ontological knowledge about domain
What determines probability?

- Statistical signals from text extractors and classifiers

- Ontological knowledge about domain
  - Functional(Spouse) & R(A,Spouse,B) -> !R(A,Spouse,C)
  - Range(Spouse, Person) & R(A,Spouse,B) -> Type(B, Person)
What determines probability?

• Statistical signals from text extractors and classifiers

• Ontological knowledge about domain

• Rules and patterns mined from data
What determines probability?

- Statistical signals from text extractors and classifiers
- Ontological knowledge about domain
- Rules and patterns mined from data
  - $R(A, \text{Spouse}, B) \& R(A, \text{Lives}, L) \rightarrow R(B, \text{Lives}, L)$
  - $R(A, \text{Spouse}, B) \& R(A, \text{Child}, C) \rightarrow R(B, \text{Child}, C)$
What determines probability?

- **Statistical signals from text extractors and classifiers**
  - \( P(R(John, Spouse, Yoko)) = 0.75; P(R(John, Spouse, Cynthia)) = 0.25 \)
  - LevenshteinSimilarity(Beatles, Beetles) = 0.9

- **Ontological knowledge about domain**
  - Functional(Spouse) & \( R(A, Spouse, B) \) \( \rightarrow \) !\( R(A, Spouse, C) \)
  - Range(Spouse, Person) & \( R(A, Spouse, B) \) \( \rightarrow \) Type(B, Person)

- **Rules and patterns mined from data**
  - \( R(A, Spouse, B) \) & \( R(A, Lives, L) \) \( \rightarrow \) R(B, Lives, L)
  - \( R(A, Spouse, B) \) & \( R(A, Child, C) \) \( \rightarrow \) R(B, Child, C)
Example: The Fab Four
Illustration of KG Identification

Uncertain Extractions:
.5: Lbl(Fab Four, novel)
.7: Lbl(Fab Four, musician)
.9: Lbl(Beatles, musician)
.8: Rel(Beatles, AlbumArtist, Abbey Road)
Illustration of KG Identification

Uncertain Extractions:
.5: Lbl(Fab Four, novel)
.7: Lbl(Fab Four, musician)
.9: Lbl(Beatles, musician)
.8: Rel(Beatles, AlbumArtist, Abbey Road)
Illustration of KG Identification

**Uncertain Extractions:**
- 5: Lbl(Fab Four, novel)
- 7: Lbl(Fab Four, musician)
- 9: Lbl(Beatles, musician)
- 8: Rel(Beatles, AlbumArtist, Abbey Road)

**Ontology:**
- Dom(albumArtist, musician)
- Mut(novel, musician)

**Extraction Graph**

PUJARA+ISWC13; PUJARA+AIMAG15
Illustration of KG Identification

**Uncertain Extractions:**
.5: Lbl(Fab Four, novel)
.7: Lbl(Fab Four, musician)
.9: Lbl(Beatles, musician)
.8: Rel(Beatles, AlbumArtist, Abbey Road)

**Ontology:**
Dom(albumArtist, musician)
Mut(novel, musician)

**Entity Resolution:**
SameEnt(Fab Four, Beatles)
Illustration of KG Identification

Uncertain Extractions:
.5: Lbl(Fab Four, novel)
.7: Lbl(Fab Four, musician)
.9: Lbl(Beatles, musician)
.8: Rel(Beatles, AlbumArtist, Abbey Road)

Ontology:
Dom(albumArtist, musician)
Mut(novel, musician)

Entity Resolution:
SameEnt(Fab Four, Beatles)

After Knowledge Graph Identification

PUJARA+ISWC13; PUJARA+AIMAG15
Probabilistic graphical model for KG

Lbl(Beatles, novel)

Lbl(Beatles, musician)

Lbl(Fab Four, novel)

Rel(Beatles, AlbumArtist, Abbey Road)

Rel(Fab Four, AlbumArtist, Abbey Road)

Lbl(Fab Four, musician)
Defining graphical models

• Many options for defining a graphical model

• We focus on two approaches, MLNs and PSL, that use rules

• MLNs treat facts as Boolean, use sampling for satisfaction

• PSL infers a “truth value” for each fact via optimization
Rules for KG Model

100: Subsumes(L1,L2) & Label(E,L1) -> Label(E,L2)
100: Exclusive(L1,L2) & Label(E,L1) -> !Label(E,L2)
100: Inverse(R1,R2) & Relation(R1,E,O) -> Relation(R2,O,E)
100: Subsumes(R1,R2) & Relation(R1,E,O) -> Relation(R2,E,O)
100: Exclusive(R1,R2) & Relation(R1,E,O) -> !Relation(R2,E,O)
100: Domain(R,L) & Relation(R,E,O) -> Label(E,L)
100: Range(R,L) & Relation(R,E,O) -> Label(O,L)
10: SameEntity(E1,E2) & Label(E1,L) -> Label(E2,L)
10: SameEntity(E1,E2) & Relation(R,E1,O) -> Relation(R,E2,O)
1: Label_OBIE(E,L) -> Label(E,L)
1: Label_OpenIE(E,L) -> Label(E,L)
1: Relation_Pattern(R,E,O) -> Relation(R,E,O)
1: !Relation(R,E,O)
1: !Label(E,L)
Rules to Distributions

• Rules are *grounded* by substituting literals into formulas

\[ w_r : \text{SAMEENT}(\text{Fab Four, Beatles}) \land \text{LBL(Beatles, musician)} \Rightarrow \text{LBL(Fab Four, musician)} \]

• Each ground rule has a *weighted satisfaction* derived from the formula’s truth value

\[
P(G|E) = \frac{1}{Z} \exp \left[ \sum_{r \in R} w_r \phi_r(G, E) \right]
\]

• Together, the ground rules provide a joint probability distribution over knowledge graph facts, conditioned on the extractions

JIANG+ICDM12; PUJARA+ISWC13
Probability Distribution over KGs

\[ P(G \mid E) = \frac{1}{Z} \exp \left[ - \sum_{r \in R} w_r \varphi_r (G) \right] \]

- \text{CANDLBL}_T(\text{FabFour, novel}) \Rightarrow \text{LBL}(\text{FabFour, novel})
- \text{MUT(}\text{novel, musician}) \Rightarrow \neg \text{LBL}(\text{Beatles, musician})
- \text{SAMEENT(}\text{Beatles, FabFour}) \Rightarrow \text{LBL}(\text{FabFour, musician})
[φ₁] \text{CANDLbl}_{\text{struct}}(\text{FabFour, novel})
    \Rightarrow \text{LBL}(\text{FabFour, novel})

[φ₂] \text{CANDRel}_{\text{pat}}(\text{Beatles, AlbumArtist, AbbeyRoad})
    \Rightarrow \text{REL}(\text{Beatles, AlbumArtist, AbbeyRoad})

[φ₃] \text{SAMEENT}(\text{Beatles, FabFour})
    \land \text{LBL}(\text{Beatles, musician})
    \Rightarrow \text{LBL}(\text{FabFour, musician})

[φ₄] \text{DOM}(\text{AlbumArtist, musician})
    \land \text{REL}(\text{Beatles, AlbumArtist, AbbeyRoad})
    \Rightarrow \text{LBL}(\text{Beatles, musician})

[φ₅] \text{MUT}(\text{musician, novel})
    \land \text{LBL}(\text{FabFour, musician})
    \Rightarrow \neg \text{LBL}(\text{FabFour, novel})
How do we get a knowledge graph?

Have: $P(KG)$ for all KGs

Need: best KG

MAP inference: optimizing over distribution to find the best knowledge graph
Inference and KG optimization

• Finding the best KG satisfying weighed rules: NP Hard

• MLNs [discrete]: Monte Carlo sampling methods
  • Solution quality dependent on burn-in time, iterations, etc.

• PSL [continuous]: optimize convex linear surrogate
  • Fast optimization, ¾-optimal MAX SAT lower bound
Graphical Models Experiments

**Data:** ~1.5M extractions, ~70K ontological relations, ~500 relation/label types

**Task:** Collectively construct a KG and evaluate on 25K target facts

**Comparisons:**
- **Extract**: Average confidences of extractors for each fact in the NELL candidates
- **Rules**: Default, rule-based heuristic strategy used by the NELL project
- **MLN**: Jiang+, ICDM12 – estimates marginal probabilities with MC-SAT
- **PSL**: Pujara+, ISWC13 – convex optimization of continuous truth values with ADMM

**Running Time:** Inference completes in 10 seconds, values for 25K facts

<table>
<thead>
<tr>
<th>Method</th>
<th>AUC</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extract</td>
<td>.873</td>
<td>.828</td>
</tr>
<tr>
<td>Rules</td>
<td>.765</td>
<td>.673</td>
</tr>
<tr>
<td>MLN (Jiang, 12)</td>
<td>.899</td>
<td>.836</td>
</tr>
<tr>
<td>PSL (Pujara, 13)</td>
<td>.904</td>
<td>.853</td>
</tr>
</tbody>
</table>

JIANG+ICDM12; PUJARA+ISWC13
Graphical Models: Pros/Cons

**BENEFITS**

- Define probability distribution over KGs
- Easily specified via rules
- Fuse knowledge from many different sources

**DRAWBACKS**

- Requires optimization over all KG facts - overkill
- Dependent on rules from ontology/expert
- Require probabilistic semantics - unavailable