Knowledge Graph Construction from Text

AAAI 2017

Jay Pujara, Sameer Singh, Bhavana Dalvi
Introducing Presenters

Jay Pujara: Postdoc at UC Santa Cruz

Sameer Singh: Assistant Professor at UCI

Bhavana Dalvi: Research Scientist at AI2
Tutorial Overview
Tutorial Overview

Part 1: Knowledge Graphs
Tutorial Overview

Part 1: Knowledge Graphs

Part 2: Knowledge Extraction
Tutorial Overview

Part 1: Knowledge Graphs

Part 2: Knowledge Extraction

Part 3: Graph Construction
Tutorial Overview

Part 1: Knowledge Graphs

Part 2: Knowledge Extraction

Part 3: Graph Construction

Part 4: Critical Analysis
Tutorial Overview

Part 1: Knowledge Graphs

Part 2: Knowledge Extraction

Part 3: Graph Construction

Part 4: Critical Analysis
Tutorial Outline

1. Knowledge Graph Primer [Jay]

2. Knowledge Extraction from Text
   a. NLP Fundamentals [Sameer]
   b. Information Extraction [Bhavana]

   Coffee Break

3. Knowledge Graph Construction
   a. Probabilistic Models [Jay]
   b. Embedding Techniques [Sameer]

4. Critical Overview and Conclusion [Bhavana]
Tutorial Overview

Part 1: Knowledge Graphs

Part 2: Knowledge Extraction

Part 3: Graph Construction

Part 4: Critical Analysis
Knowledge Graph Primer

TOPICS:

What is a Knowledge Graph?
Why are Knowledge Graphs Important?
Where do Knowledge Graphs come from?
Knowledge Representation Choices
Problem Overview
Knowledge Graph Primer

TOPICS:

**What is a Knowledge Graph?**

**Why are Knowledge Graphs Important?**

**Where do Knowledge Graphs come from?**

**Knowledge Representation Choices**

**Problem Overview**
What is a knowledge graph?
What is a knowledge graph?

• Knowledge in graph form!
What is a knowledge graph?

• Knowledge in graph form!

• Captures entities, attributes, and relationships
What is a knowledge graph?

- Knowledge in graph form!
- Captures entities, attributes, and relationships
- Nodes are entities
What is a knowledge graph?

• Knowledge in graph form!

• Captures entities, attributes, and relationships

• Nodes are entities

• Nodes are labeled with attributes (e.g., types)
What is a knowledge graph?

- Knowledge in graph form!
- Captures entities, attributes, and relationships
- Nodes are entities
- Nodes are labeled with attributes (e.g., types)
- Typed edges between two nodes capture a relationship between entities
Example knowledge graph

- Knowledge in graph form!
- Captures entities, attributes, and relationships
  - Nodes are entities
  - Nodes are labeled with attributes (e.g., types)
  - Typed edges between two nodes capture a relationship between entities
Knowledge Graph Primer

TOPICS:

What is a Knowledge Graph?

Why are Knowledge Graphs Important?

Where do Knowledge Graphs come from?

Knowledge Representation Choices

Problem Overview
Why knowledge graphs?

• Humans:
  • Combat information overload
  • Explore via intuitive structure
  • Tool for supporting knowledge-driven tasks

• AIs:
  • Key ingredient for many AI tasks
  • Bridge from data to human semantics
  • Use decades of work on graph analysis
Applications 1: QA/Agents

"What is the weather at the Hilton in San Francisco?"
OK, here's the weather for the Hilton for today:

San Francisco
Partly Cloudy

53°

Monday Today
12 PM 55
4 PM 53

who is playing in this year's super bowl

Super Bowl LI
New England Patriots
Atlanta Falcons
Tickets - Preview

About 15,300,000 results (0.66 seconds)
All times are in Eastern Time
Applications 2: Decision Support
Applications 3: Fueling Discovery

**beatles (musicartist)**

literal strings: **BEATLES, Beatles, beatles**

---

Help NELL Learn!

NELL wants to know if these be
If they are or ever were, click thumbs-up. O

- beatles is a musical artist
- beatles is a musician in the genre classic pop (musicgenre)
- beatles is a musician in the genre pop (musicgenre)
- beatles is a musician in the genre rock (musicgenre)
- beatles is a musician in the genre classic rock (musicgenre)
Knowledge Graphs & Industry

- Google Knowledge Graph
  - Google Knowledge Vault
- Amazon Product Graph
- Facebook Graph API
- IBM Watson
- Microsoft Satori
  - Project Hanover/Literome
- LinkedIn Knowledge Graph
- Yandex Object Answer
- Diffbot, GraphIQ, Maana, ParseHub, Reactor Labs, SpazioDati
Knowledge Graph Primer

TOPICS:

What is a Knowledge Graph?
Why are Knowledge Graphs Important?
Where do Knowledge Graphs come from?
Knowledge Representation Choices
Problem Overview
Where do knowledge graphs come from?
Where do knowledge graphs come from?

- **Structured Text**
  - Wikipedia Infoboxes, tables, databases, social nets
Where do knowledge graphs come from?

- **Structured Text**
  - Wikipedia Infoboxes, tables, databases, social nets

- **Unstructured Text**
  - WWW, news, social media, reference articles

---

**Beatles last live performance**

Published: Thursday, January 26th 2017, 5:24 am PST
Updated: Monday, January 30th 2017, 4:06 am PST
Written by Jim Ettink, Producer

(KFVS) - How about a little Beatles history.

It was on this date in 1969, the band performed their last live public performance.

*Allan Williams, First Manager of the Beatles, Dies at 86*
Where do knowledge graphs come from?

• Structured Text
  ◦ Wikipedia Infoboxes, tables, databases, social nets

• Unstructured Text
  ◦ WWW, news, social media, reference articles

• Images
Where do knowledge graphs come from?

- **Structured Text**
  - Wikipedia Infoboxes, tables, databases, social nets

- **Unstructured Text**
  - WWW, news, social media, reference articles

- **Images**

- **Video**
  - YouTube, video feeds
Knowledge Graph Primer

TOPICS:

What is a Knowledge Graph?
Why are Knowledge Graphs Important?
Where do Knowledge Graphs come from?
Knowledge Representation Choices
Problem Overview
Knowledge Representation

- Decades of research into knowledge representation

- Most knowledge graph implementations use RDF triples
  - `<rdf:subject, rdf:predicate, rdf:object> : r(s,p,o)

- Temporal scoping, reification, and skolemization...

- ABox (assertions) versus TBox (terminology)

- Common ontological primitives
  - rdfs:domain, rdfs:range, rdf:type, rdfs:subClassOf, rdfs:subPropertyOf, ...
  - owl:inverseOf, owl:TransitiveProperty, owl:FunctionalProperty, ...
Semantic Web

- Standards for defining and exchanging knowledge
  - RDF, RDFa, JSON-LD, schema.org
  - RDFS, OWL, SKOS, FOAF

- Annotated data provide critical resource for automation

- Major weakness: annotate everything?

"LINKING OPEN DATA CLOUD DIAGRAM 2014, BY MAX SCHMACHTENBERG, CHRISTIAN BIZER, ANJA JENTZSCH AND RICHARD CYGANKI. HTTP://LOD-CLOUD.NET/"
Information Extraction from Text

• Focus of this tutorial!

• Answer to the knowledge acquisition bottleneck

• Many challenges:
  • chunking
  • polysemy/word sense disambiguation
  • entity coreference
  • relational extraction
Knowledge Graph Primer

TOPICS:

What is a Knowledge Graph?
Why are Knowledge Graphs Important?
Where do Knowledge Graphs come from?
Knowledge Representation Choices

Problem Overview
What is a knowledge graph?

- Knowledge in graph form!

- Captures entities, attributes, and relationships

- Nodes are entities

- Nodes are labeled with attributes (e.g., types)

- Typed edges between two nodes capture a relationship between entities
Basic problems
Basic problems

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

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

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

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

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

- **How** are they related (edges)?
Basic problems

- **Who** are the entities (nodes) in the graph?
- **What** are their attributes and types (labels)?
- **How** are they related (edges)?
Knowledge Graph Construction
Two perspectives

Knowledge Extraction

• **Who** are the entities (nodes) in the graph?
  • Named Entity Recognition
  • Entity Coreference

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

• **How** are they related (edges)?
  • Relation Extraction
  • Semantic Role Labeling

Graph Construction

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

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

• **How** are they related (edges)?
  • Link Prediction
Tutorial Outline

1. Knowledge Graph Primer [Jay]

2. Knowledge Extraction from Text
   a. NLP Fundamentals [Sameer]
   b. Information Extraction [Bhavana]

Coffee Break

3. Knowledge Graph Construction
   a. Probabilistic Models [Jay]
   b. Embedding Techniques [Sameer]

4. Critical Overview and Conclusion [Bhavana]