Knowledge Base
I can recognize everything in the room (proudly)

Bring me a cup of hot water!

Well, I can tell you “where is the cup?”

⇒ Recognize everything, but can do nothing
What is missing?

Bring me a cup of hot water

• find a cup
• realize a cup has *containable* affordance
Affordance

- A cup
- grasp
- filled in water
- pour

Attribute

- A cup
- brittle
- made of glass, plastic
- has a handle
What is missing?

Bring me a cup of hot water

• find a cup
• realize a cup has *containable* affordance
• cup is empty
• find tape, fill in water
• find microwave
• heat it up

The Common Knowledge
The Common Knowledge
Structured

Specific

General

Casual format

Structured

Specific

General

Casual format

WordNet
A lexical database for English

SUN

IMAGENET

Google

WIKIPEDIA
The Free Encyclopedia

eBay

Merriam-Webster's Collegiate Dictionary

Amazon

Updated Annually

The Oxford Encyclopedia of American Military & Diplomatic History

Casual format

Structured
DBpedia

DBpedia is a crowd-sourced community effort to extract structured information from Wikipedia.
First-order logic

From Wikipedia, the free encyclopedia

First-order logic is a formal system used in mathematics, philosophy, linguistics, and computer science. It is also known as first-order predicate calculus, the lower predicate calculus, quantification theory, and predicate logic. First-order logic uses quantified variables over (non-logical) objects. This distinguishes it from propositional logic which does not use quantifiers.

A theory about some topic is usually first-order logic together with a specified domain of discourse over which the quantified variables range, finitely many functions which map from that domain into itself, finitely many predicates defined on that domain, and a recursive set of axioms which are believed to hold for those things. Sometimes theory is understood in a more formal sense, which is just a set of sentences in first-order logic.

The adjective “first-order” distinguishes first-order logic from higher-order logic in which there are predicates having predicates or functions as arguments, or in which one or both of predicate quantifiers or function quantifiers are permitted. In first-order theories, predicates are often associated with sets. In interpreted higher-order theories, predicates may be interpreted as sets of sets.

There are many deductive systems for first-order logic that are sound (all provable statements are true in all models) and complete (all statements which are true in all models are provable). Although the logical consequence relation is only semidecidable, much progress has been made in automated theorem proving in first-order logic. First-order logic also satisfies several metalogical theorems that make it amenable to analysis in proof theory, such as the Löwenheim–Skolem theorem and the compactness theorem.

First-order logic is the standard for the formalization of mathematics into axioms and is studied in the foundations of mathematics. Mathematical theories, such as number theory and set theory, have been formalized as first-order theories, and proofs in such theories are routines that manipulate the symbols of the theory, according to the axioms and the rules of inference.

For a history of first-order logic and how it came to dominate formal logic, see José Ferreirós (2001).

Contents

1 Introduction
2 Syntax
  2.1 Alphabet
  2.1.1 Logical symbols
  2.1.2 Non-logical symbols
3 Formation rules
4 Axioms
  4.1 Axioms in set theory
  4.2 Other axioms
5 Semantics
  5.1 First-order structures
  5.2 Evaluation of truth values
  5.3 Validity, satisfiability, and logical consequence
  5.4 Axiomatizations
6 First-order theories, models, and elementary classes
7 Empty domains

About: First-order logic

An Entity of Type "Logic" from NAMED_GRAPH

First-order logic is a formal system used in mathematics, philosophy, linguistics, and computer science. It is also known as first-order predicate calculus, the lower predicate calculus, quantification theory, and predicate logic.

Properties

- First-order logic is a formal system used in mathematics, philosophy, linguistics, and computer science. It is also known as first-order predicate calculus, the lower predicate calculus, quantification theory, and predicate logic. First-order logic is the standard for the formalization of mathematics into axioms and is studied in the foundations of mathematics. Mathematical theories, such as number theory and set theory, have been formalized as first-order theories, and proofs in such theories are routines that manipulate the symbols of the theory, according to the axioms and the rules of inference.

- First-order logic is a formal system used in mathematics, philosophy, linguistics, and computer science. It is also known as first-order predicate calculus, the lower predicate calculus, quantification theory, and predicate logic. First-order logic is the standard for the formalization of mathematics into axioms and is studied in the foundations of mathematics. Mathematical theories, such as number theory and set theory, have been formalized as first-order theories, and proofs in such theories are routines that manipulate the symbols of the theory, according to the axioms and the rules of inference.

- First-order logic is a formal system used in mathematics, philosophy, linguistics, and computer science. It is also known as first-order predicate calculus, the lower predicate calculus, quantification theory, and predicate logic. First-order logic is the standard for the formalization of mathematics into axioms and is studied in the foundations of mathematics. Mathematical theories, such as number theory and set theory, have been formalized as first-order theories, and proofs in such theories are routines that manipulate the symbols of the theory, according to the axioms and the rules of inference.

- First-order logic is a formal system used in mathematics, philosophy, linguistics, and computer science. It is also known as first-order predicate calculus, the lower predicate calculus, quantification theory, and predicate logic. First-order logic is the standard for the formalization of mathematics into axioms and is studied in the foundations of mathematics. Mathematical theories, such as number theory and set theory, have been formalized as first-order theories, and proofs in such theories are routines that manipulate the symbols of the theory, according to the axioms and the rules of inference.

- First-order logic is a formal system used in mathematics, philosophy, linguistics, and computer science. It is also known as first-order predicate calculus, the lower predicate calculus, quantification theory, and predicate logic. First-order logic is the standard for the formalization of mathematics into axioms and is studied in the foundations of mathematics. Mathematical theories, such as number theory and set theory, have been formalized as first-order theories, and proofs in such theories are routines that manipulate the symbols of the theory, according to the axioms and the rules of inference.

- First-order logic is a formal system used in mathematics, philosophy, linguistics, and computer science. It is also known as first-order predicate calculus, the lower predicate calculus, quantification theory, and predicate logic. First-order logic is the standard for the formalization of mathematics into axioms and is studied in the foundations of mathematics. Mathematical theories, such as number theory and set theory, have been formalized as first-order theories, and proofs in such theories are routines that manipulate the symbols of the theory, according to the axioms and the rules of inference.

- First-order logic is a formal system used in mathematics, philosophy, linguistics, and computer science. It is also known as first-order predicate calculus, the lower predicate calculus, quantification theory, and predicate logic. First-order logic is the standard for the formalization of mathematics into axioms and is studied in the foundations of mathematics. Mathematical theories, such as number theory and set theory, have been formalized as first-order theories, and proofs in such theories are routines that manipulate the symbols of the theory, according to the axioms and the rules of inference.
A general method for conceptual description or modeling of information that is implemented in web resources.

Make statements about web resources in the form of subject-predicate-object expression.
There is a Person identified by http://www.w3.org/People/EM/contact#me, whose name is Eric Miller, whose email address is e.miller123(at)example (changed for security purposes), and whose title is Dr.

- Subject: "http://www.w3.org/People/EM/contact#me"
- The objects are:
  - "Eric Miller" (with a predicate "whose name is"),
  -mailto:e.miller123(at)example (with a predicate "whose email address is"),
  and
  - "Dr." (with a predicate "whose title is").
- The predicates also have URIs. For example, the URI for each predicate:
  - "whose name is" is http://www.w3.org/2000/10/swap/pim/contact#fullName,
  - "whose email address is" is http://www.w3.org/2000/10/swap/pim/contact#mailbox,
  - "whose title is" is http://www.w3.org/2000/10/swap/pim/contact#personalTitle.

-RDF triples can be expressed:

- http://www.w3.org/People/EM/contact#me, http://www.w3.org/2000/10/swap/pim/contact#fullName, "Eric Miller"
- http://www.w3.org/People/EM/contact#me, http://www.w3.org/2000/10/swap/pim/contact#mailbox, mailto:e.miller123(at)example
- http://www.w3.org/People/EM/contact#me, http://www.w3.org/2000/10/swap/pim/contact#personalTitle, "Dr."
- http://www.w3.org/People/EM/contact#me, http://www.w3.org/1999/02/22-rdf-syntax-ns#type, http://www.w3.org/2000/10/swap/pim/contact#Person
DBpedia

➤ Revolutionize Wikipedia Search

➤ “Tell me all the episodes of Game of Thrones” rank them by released date.

```sql
SELECT *
WHERE {
} ORDER BY DESC(?date)
```
<table>
<thead>
<tr>
<th></th>
<th>date</th>
<th>number</th>
<th>season</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://dbpedia.org/resource/Oathkeeper">http://dbpedia.org/resource/Oathkeeper</a></td>
<td>2014-04-27+02:00</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td><a href="http://dbpedia.org/resource/Breaker_of_Chains">http://dbpedia.org/resource/Breaker_of_Chains</a></td>
<td>2014-04-20+02:00</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td><a href="http://dbpedia.org/resource/The_Lion_and_the_Rose">http://dbpedia.org/resource/The_Lion_and_the_Rose</a></td>
<td>2014-04-13+02:00</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td><a href="http://dbpedia.org/resource/Two_Swords_(Game_of_Thrones)">http://dbpedia.org/resource/Two_Swords_(Game_of_Thrones)</a></td>
<td>2014-04-06+02:00</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td><a href="http://dbpedia.org/resource/Mhysa">http://dbpedia.org/resource/Mhysa</a></td>
<td>2013-06-09+02:00</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td><a href="http://dbpedia.org/resource/The_Rains_of_Castamere">http://dbpedia.org/resource/The_Rains_of_Castamere</a></td>
<td>2013-06-02+02:00</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td><a href="http://dbpedia.org/resource/Second_Sons">http://dbpedia.org/resource/Second_Sons</a></td>
<td>2013-05-19+02:00</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td><a href="http://dbpedia.org/resource/The_Bear_and_the_Maiden_Fair">http://dbpedia.org/resource/The_Bear_and_the_Maiden_Fair</a></td>
<td>2013-05-12+02:00</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td><a href="http://dbpedia.org/resource/The_Climb_(Game_of_Thrones)">http://dbpedia.org/resource/The_Climb_(Game_of_Thrones)</a></td>
<td>2013-05-05+02:00</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td><a href="http://dbpedia.org/resource/Kissed_by_Fire">http://dbpedia.org/resource/Kissed_by_Fire</a></td>
<td>2013-04-28+02:00</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td><a href="http://dbpedia.org/resource/And_Now_His_Watch_Is_Ended">http://dbpedia.org/resource/And_Now_His_Watch_Is_Ended</a></td>
<td>2013-04-21+02:00</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td><a href="http://dbpedia.org/resource/Walk_of_Punishment">http://dbpedia.org/resource/Walk_of_Punishment</a></td>
<td>2013-04-14+02:00</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td><a href="http://dbpedia.org/resource/Dark_Wings,_Dark_Words">http://dbpedia.org/resource/Dark_Wings,_Dark_Words</a></td>
<td>2013-04-07+02:00</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td><a href="http://dbpedia.org/resource/Valar_Dohaeris">http://dbpedia.org/resource/Valar_Dohaeris</a></td>
<td>2013-03-31+02:00</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td><a href="http://dbpedia.org/resource/Valar_Morghulis">http://dbpedia.org/resource/Valar_Morghulis</a></td>
<td>2012-06-03+02:00</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td><a href="http://dbpedia.org/resource/Blackwater_(Game_of_Thrones)">http://dbpedia.org/resource/Blackwater_(Game_of_Thrones)</a></td>
<td>2012-05-27+02:00</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td><a href="http://dbpedia.org/resource/The_Prince_of_Winterfell">http://dbpedia.org/resource/The_Prince_of_Winterfell</a></td>
<td>2012-05-20+02:00</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td><a href="http://dbpedia.org/resource/A_Man_Without_Honor">http://dbpedia.org/resource/A_Man_Without_Honor</a></td>
<td>2012-05-13+02:00</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td><a href="http://dbpedia.org/resource/The_Old_Gods_and_the_New">http://dbpedia.org/resource/The_Old_Gods_and_the_New</a></td>
<td>2012-05-06+02:00</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td><a href="http://dbpedia.org/resource/The_Ghost_of_Harrenhal">http://dbpedia.org/resource/The_Ghost_of_Harrenhal</a></td>
<td>2012-04-29+02:00</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td><a href="http://dbpedia.org/resource/Garden_of_Bones">http://dbpedia.org/resource/Garden_of_Bones</a></td>
<td>2012-04-22+02:00</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td><a href="http://dbpedia.org/resource/What_Is_Dead_May_Never_Die">http://dbpedia.org/resource/What_Is_Dead_May_Never_Die</a></td>
<td>2012-04-15+02:00</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td><a href="http://dbpedia.org/resource/The_Night_Lands">http://dbpedia.org/resource/The_Night_Lands</a></td>
<td>2012-04-08+02:00</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td><a href="http://dbpedia.org/resource/The_North_Remembers">http://dbpedia.org/resource/The_North_Remembers</a></td>
<td>2012-04-01+02:00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td><a href="http://dbpedia.org/resource/Fire_and_Blood_(Game_of_Thrones)">http://dbpedia.org/resource/Fire_and_Blood_(Game_of_Thrones)</a></td>
<td>2011-06-19+02:00</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td><a href="http://dbpedia.org/resource/Baelor">http://dbpedia.org/resource/Baelor</a></td>
<td>2011-06-12+02:00</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td><a href="http://dbpedia.org/resource/The_Pointy_End">http://dbpedia.org/resource/The_Pointy_End</a></td>
<td>2011-06-05+02:00</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td><a href="http://dbpedia.org/resource/You_Win_or_You_Die">http://dbpedia.org/resource/You_Win_or_You_Die</a></td>
<td>2011-05-29+02:00</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td><a href="http://dbpedia.org/resource/A_Golden_Crown">http://dbpedia.org/resource/A_Golden_Crown</a></td>
<td>2011-05-22+02:00</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td><a href="http://dbpedia.org/resource/The_Wolf_and_the_Lion">http://dbpedia.org/resource/The_Wolf_and_the_Lion</a></td>
<td>2011-05-15+02:00</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td><a href="http://dbpedia.org/resource/Cripples">http://dbpedia.org/resource/Cripples</a>,_ Bastards, and Broken Things</td>
<td>2011-05-08+02:00</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td><a href="http://dbpedia.org/resource/Lord_Snow">http://dbpedia.org/resource/Lord_Snow</a></td>
<td>2011-05-01+02:00</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td><a href="http://dbpedia.org/resource/The_Kingsroad">http://dbpedia.org/resource/The_Kingsroad</a></td>
<td>2011-04-24+02:00</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td><a href="http://dbpedia.org/resource/Winter_Is_Coming">http://dbpedia.org/resource/Winter_Is_Coming</a></td>
<td>2011-04-17+02:00</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
DBpedia

- A lot of other applications
- [http://wiki.dbpedia.org/Applications](http://wiki.dbpedia.org/Applications)
- Available in multiple languages
- Downloadable
Knowledge Base

- Source of knowledge: internet, human input
- Structure: Graph = Node + Edge
  - RDF: subject-predicate-object
  - Node: entity
  - Edge: relation
WikiData

• Very similar as DBpedia
• link to more source
• act as knowledge base for Wikimedia
Wait, wait...

- Knowledge base, structured data organized in graph, DBpedia, Wikidata, Freebase.

- But...

  Bring me a cup of hot water

- Need low level knowledge
  - find a cup
  - a cup has containable affordance
  - cup is empty
  - find tape, fill in water
  - find microwave
  - heat it up
ConceptNet

- A semantic network containing lots of things computers should know about the world.

ConceptNet contains everyday basic knowledge:

learn — MotivatedByGoal → knowledge
You would learn because you want knowledge.

Cultural knowledge:

saxophone — UsedFor → jazz
A saxophone is used for jazz.

And scientific knowledge:

semantic role — HasContext → linguistics
"Semantic role" is a term in linguistics.

It would not adequately represent human knowledge if it didn’t contain other languages besides English, as well:

本 — MadeOf → 紙
本は紙でできている。(A book is made of paper.)
ConceptNet

- Free to download
- Provide API to:
  - Retrieve the data for particular nodes and edges
  - Query for edges with given properties
  - Measure and query the semantic distance between nodes
So far...

- There are lexical knowledge base for both high-level and low-level knowledge ready online.
- To connect the knowledge with computer vision, we need visual knowledge base.
- Not as explicit as language
- “A car can be used for driving”
Never Ending Image Learner

- Learn from image searching engine (the weak association between image and text)
- what a car looks like?
- know that sheep are white
Never Ending Image Learner

- NEIL is a computer program
- Run 24h per day, 7 days per week
- Automatically extract visual knowledge from internet data
- Learn to see
- Learn common sense

Airbus_330 can be a kind of / look similar to Airplane.
Bamboo_forest can be / can have Vertical_lines.
Car can have a part Wheel.
Trading_floor can be / can have Crowded.
Never Ending Image Learner

(0) Seed Images

(1) Visual Cluster Discovery (Section 3.1)

(2) Train Detectors

(3) Relationship Discovery (Section 3.2)

(4) Add New Instances (Section 3.3)

(5) Retrain Detectors

Learned facts:
- Monitor is a part of Desktop Computer
- Keyboard is a part of Desktop Computer
- Television looks similar to Monitor
Never Ending Image Learner

- Seeding Classifier via Google Image Search
  - scene, attribute classifier; object, attribute detector.
  - Directly train scene and attribute classifier on downloaded images.
- However, fail for object and attribute detector
  - Outlier, Polysemy, Visual diversity, Localization
Never Ending Image Learner

- Seeding Classifier via Google Image Search
- Train exemplar-LDA for each image
- Run detection on all images
- Get top K windows with high scores from multiple detectors
- Clustering with ELDA score vector
- Train classifier for each cluster
Never Ending Image Learner

Seeding Classifier via Google Image Search

(a) Google Image Search for “tricycle”

(b) Sub-category Discovery
Never Ending Image Learner

- Extract Relationships

- Object-Object Relationships:
  - Partonomy: Eye is a part of Baby.
  - Taxonomy: BMW 320 is a kind of Car.
  - Similarity: Swan looks similar to Goose.
Never Ending Image Learner

- Extract Relationships
  - Build co-occurrence matrix
  - Get co-occurred object pairs
  - Learn relationship in terms of mean and variance of relative positive, aspect ratio, score, size.
Never Ending Image Learner

- Object-Attribute Relationships
  - “Pizza has Round Shape”, “Sunflower is Yellow”
- Scene-Object Relationships
  - “Bus is found in Bus depot”
- Scene-Attribute Relationships
  - “Ocean is Blue”
Never Ending Image Learner

- Discover new instance and retrain

\[
\phi_i(x) + \sum_{i,j \in \mathcal{R}_O \cup \mathcal{R}_A} \phi_j(x_l) \psi_{i,j}(x, x_l) + \sum_{i,k \in \mathcal{R}_S} \omega_k(x)
\]

object detector  binary relationship

all related objects and attributes  scene classifier

all related scenes
Never Ending Image Learner

(0) Seed Images

(1) Visual Cluster Discovery (Section 3.1)

(1) Desktop Computer
(2) Monitor
(3) Keyboard
(4) Television

(2) Train Detectors

(3) Relationship Discovery (Section 3.2)

Desktop Computer (1)
Desktop Computer (2)
Desktop Computer (3)
Monitor (1)

(4) Add New Instances (Section 3.3)

(5) Retrain Detectors

Learned facts:
- Monitor is a part of Desktop Computer
- Keyboard is a part of Desktop Computer
- Television looks similar to Monitor
Never Ending Image Learner

- **Bootstrapping**
- **Words:** NELL (never ending language learning)
- **Images:** ImageNet, SUN, Google Image Search

<table>
<thead>
<tr>
<th>Table 1. mAP performance for scene classification on 12 categories.</th>
<th>mAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seed Classifier (15 Google Images)</td>
<td>0.52</td>
</tr>
<tr>
<td>Bootstrapping (without relationships)</td>
<td>0.54</td>
</tr>
<tr>
<td>NEIL Scene Classifiers</td>
<td>0.57</td>
</tr>
<tr>
<td>NEIL (Classifiers + Relationships)</td>
<td><strong>0.62</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2. mAP performance for object detection on 15 categories.</th>
<th>mAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latent SVM (50 Google Images)</td>
<td>0.34</td>
</tr>
<tr>
<td>Latent SVM (450 Google Images)</td>
<td>0.28</td>
</tr>
<tr>
<td>Latent SVM (450, Aspect Ratio Clustering)</td>
<td>0.30</td>
</tr>
<tr>
<td>Latent SVM (450, HOG-based Clustering)</td>
<td>0.33</td>
</tr>
<tr>
<td>Seed Detector (NEIL Clustering)</td>
<td>0.44</td>
</tr>
<tr>
<td>Bootstrapping (without relationships)</td>
<td>0.45</td>
</tr>
<tr>
<td>NEIL Detector</td>
<td>0.49</td>
</tr>
<tr>
<td>NEIL Detector + Relationships</td>
<td><strong>0.51</strong></td>
</tr>
</tbody>
</table>
Hey, it’s about time…

- to fix the annoying problem

Bring me a cup of hot water

- Design a robot with knowledge base
RoboBrain

- A large-scale knowledge engine for robot
- Build a knowledge base similar as ConceptNet
- More diverse edges
- Edges have beliefs
  - measure the confidence of learned relations
  - labelled by crowd-sourced feedback
RoboBrain
How to build knowledge base?

again, graph represented in triplets

<table>
<thead>
<tr>
<th>Word</th>
<th>An english word represented as an ASCII string.</th>
</tr>
</thead>
<tbody>
<tr>
<td>DeepFeature</td>
<td>Feature function trained with a Deep Neural Network</td>
</tr>
<tr>
<td>Image</td>
<td>2D RGB Image.</td>
</tr>
<tr>
<td>PointCloud</td>
<td>3D point cloud</td>
</tr>
<tr>
<td>IsTypeOf</td>
<td>human <em>IsTypeOf</em> a mammal.</td>
</tr>
<tr>
<td>HasAppearance</td>
<td>floor <em>HasAppearance</em> as follows (this image).</td>
</tr>
<tr>
<td>CanPerformAction</td>
<td>human <em>CanPerformAction</em> cutting.</td>
</tr>
<tr>
<td>SpatiallyDistributedAs</td>
<td>location of human is <em>SpatiallyDistributedAs</em>.</td>
</tr>
</tbody>
</table>

(StandingHuman, Shoe, *CanUse*)
(Grasping, DeepFeature23, *UsesFeature*)
(StandingHuman, $\mathcal{N}(\mu, \Sigma)$, *SpatiallyDistributedAs*)
RoboBrain

Knowledge acquisition

$G = (V, E)$

$+ (v^1_1, v^1_2, \ell^1) \ldots (v^N_1, v^N_2, \ell^N)$

$\rightarrow G' = (V', E')$

$V' = v^1_1 \cup v^1_2 \cup \ldots \cup v^N_1 \cup v^N_2 \cup V$

$E' = (v^1_1, v^1_2, \ell^1) \cup \ldots \cup (v^N_1, v^N_2, \ell^N) \cup E$
RoboBrain

Merge and Split

(a) original graph

(b) feed insertion

(c) after merge(Mug, Mug') → Mug ∘ split(Cup) → (Cup, Mug')
RoboBrain

Visualization of Knowledge Base

50K nodes, 100K edges
RoboBrain

Grounding a natural language sentence

“fill a cup with water”

moveto(cup_{01}); grasp(cup_{01}); moveto(sink_{01}); keep(cup_{01}, on, sink_{01}); toggle(sink_knob_{01}); wait(); toggle(sink_knob_{01});
RoboBrain

- Grounding a natural language sentence
- appearance, affordance, possible action, associated trajectory, manipulation feature

“add ice-cream to cup and drizzle syrup over it”
RoboBrain

☞ Support action planning

\[
\text{squeeze}(\text{syrup}_1, \text{pot}_1)
\]

- \textit{grasping pr2 syrup}_1: Robot is grasping the syrup.
- \textit{squeezeable syrup}_1: Syrup bottle should be squeezeable.
- \textit{on syrup}_1 \textit{pot}_1: Syrup bottle should be placed directly above the pot.

\[
\text{squeezeable syrup}_1 = \text{len fetch (u\{name: 'syrup'\})} \rightarrow \left[\text{HasAffordance'}\right] \rightarrow (v\{name: 'squeezeable'\}) > 0
\]
RoboBrain

转让动作原语到轨迹

\[ pour \ cup_{01} = \text{fetch} \left( \{ \text{name} : 'pour' \} \right) \rightarrow \]
\[ (v\{ \text{name} : 'HasTrajectory' \}) \leftarrow (\{ \text{name} : 'cup' \}) \]
RoboBrain

- Other application
- anticipating human activity
RoboBrain

➤ Summary

➤ a knowledge base integrates knowledge about physical world that robots live in.

➤ share knowledge to support complicated tasks

➤ natural language grounding

➤ activity prediction
Can we do more?

✧ So far, we know how to reuse learned knowledge.

✧ Can we generalize the learned knowledge to understand what we never seen before?

edible
Zero-shot Affordance Prediction

- Idea

- affordance, attribute, human interaction are highly correlated
Zero-shot Affordance Prediction

- Learning the knowledge base: choose 40 objects (Stanford 40 Action Database)

- Nodes (Entities):
  - Attribute:
    - visual: 33 per-trained classifiers, “round”, “shiny”
    - physical: weight, size, from FreeBase, Amazon
    - categorical: 22 from WordNet, “animal”, “vehicle”
Zero-shot Affordance Prediction

.Nodes

.Attributes

.Affordance

.choose 14 from Stanford 40 Action

.manual labeling for 40 objects

.on average, 4.25 per object
Zero-shot Affordance Prediction

 Nodes:

 Human pose: cluster centroids of descriptor.

 Human object relative position

(a) pose descriptor  

(b) relative locations
Zero-shot Affordance Prediction

- Learn a Markov Logic Network (MRF) to represent the relationships between nodes
- Use training data to build such relationships

![Diagram](image-url)
Zero-shot Affordance Prediction

- Zero-shot prediction:
  - choose 22 objects that are semantically similar as the 40 training objects.
  - sample 50 images per objects as testing set.
Zero-shot Affordance Prediction

- Zero-shot prediction:
  - Estimating visual attributes: run classifiers

- Inferring:
  - Categorical attributes: learn regression from image feature and VA
  - Physical attributes: regression from image feature
Zero-shot Affordance Prediction

Zero-shot prediction:

Now, we have confidence on attribute nodes. Run belief propagation on MRF, we get confidence on affordance nodes.
Zero-shot Affordance Prediction

Zero-shot prediction:
Zero-shot Affordance Prediction

Zero-shot prediction:

**Table 1.** Performance of Zero-shot Affordance Prediction (measured in mAUC)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>base features (BF)</td>
<td>0.7858</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>visual attributes (VA)</td>
<td>0.7525</td>
<td>0.7533</td>
<td>0.7432</td>
</tr>
<tr>
<td>categorical &amp; physical (CP)</td>
<td>0.7919</td>
<td>0.7924</td>
<td>0.8234</td>
</tr>
<tr>
<td>combined (VA+CP)</td>
<td>0.8006</td>
<td>0.7985</td>
<td>0.8409</td>
</tr>
</tbody>
</table>

**Table 2.** Performance of Estimating Human Poses (in Hamming distance)

<table>
<thead>
<tr>
<th>Method</th>
<th>nearest neighbor</th>
<th>attributes</th>
<th>affordances</th>
<th>attributes + affordances</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance</td>
<td>0.928</td>
<td>1.027</td>
<td>0.630</td>
<td>0.527</td>
</tr>
</tbody>
</table>
Zero-shot Affordance Prediction

Prediction from human pose:

Table 3. Predicting Actions and Objects from Human-Object Interactions

<table>
<thead>
<tr>
<th>Method</th>
<th>Action</th>
<th>Object</th>
</tr>
</thead>
<tbody>
<tr>
<td>human poses</td>
<td>50.4%</td>
<td>46.2%</td>
</tr>
<tr>
<td>poses + locations</td>
<td>81.2%</td>
<td>64.5%</td>
</tr>
</tbody>
</table>
Zero-shot Affordance Prediction

Robust to partial observation:

Fig. 12: **Performance variations against partial observation.** The $x$-axis denotes the percentage of unobserved evidence. The $y$-axis denotes the performance (mAUC). The top two curves correspond to our method. The bottom two are the classification-based method. In comparison, the knowledge base representation is more robust against partial observation.
Zero-shot Affordance Prediction

Question Answering:

<table>
<thead>
<tr>
<th>Question</th>
<th>Evidence</th>
<th>Query</th>
<th>Top Answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>What do animals look like?</td>
<td>isA(N1, Animal)</td>
<td>hasVisualAttribute(N1, x)</td>
<td>hasVisualAttribute(N1, Leather)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>hasVisualAttribute(N1, Head)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>hasVisualAttribute(N1, Tail)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>hasVisualAttribute(N1, Furry)</td>
</tr>
<tr>
<td>I saw something shiny and metallic. What is it?</td>
<td>hasVisualAttribute(N1, Shiny)</td>
<td>isA(N1, x)</td>
<td>isA(N1, Instrumentality)</td>
</tr>
<tr>
<td></td>
<td>hasVisualAttribute(N1, Metal)</td>
<td></td>
<td>isA(N1, Device)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>isA(N1, Container)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>isA(N1, Computer)</td>
</tr>
<tr>
<td>Here is a vehicle and it’s quite heavy. What can I do with it?</td>
<td>isA(N1, Vehicle)</td>
<td>hasAffordance(N1, x)</td>
<td>hasAffordance(N1, Ride)</td>
</tr>
<tr>
<td></td>
<td>hasWeight(N1, W4) (&gt; 100 kg)</td>
<td></td>
<td>hasAffordance(N1, Row)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>hasAffordance(N1, SitOn)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>hasAffordance(N1, Fix)</td>
</tr>
<tr>
<td>Tell me how heavy and large a wooden musical instrument is.</td>
<td>isA(N1, Musical_instrument)</td>
<td>hasWeight(N1, x)</td>
<td>hasSize(N1, D2) (10-100 in)</td>
</tr>
<tr>
<td></td>
<td>hasVisualAttribute(N1, Wood)</td>
<td></td>
<td>hasWeight(N1, W2) (1-10 kg)</td>
</tr>
<tr>
<td></td>
<td>hasSize(N1, x)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Summary

- Online knowledge base
  - high-level: DBpedia, Wikidata
  - low-level: ConceptNet
- How to learn visual knowledge base: NEIL
- How to create KB for robot to do complicated tasks: RoboBrain
- How to generalize KB: zero-shot affordance prediction