Symbolic AI

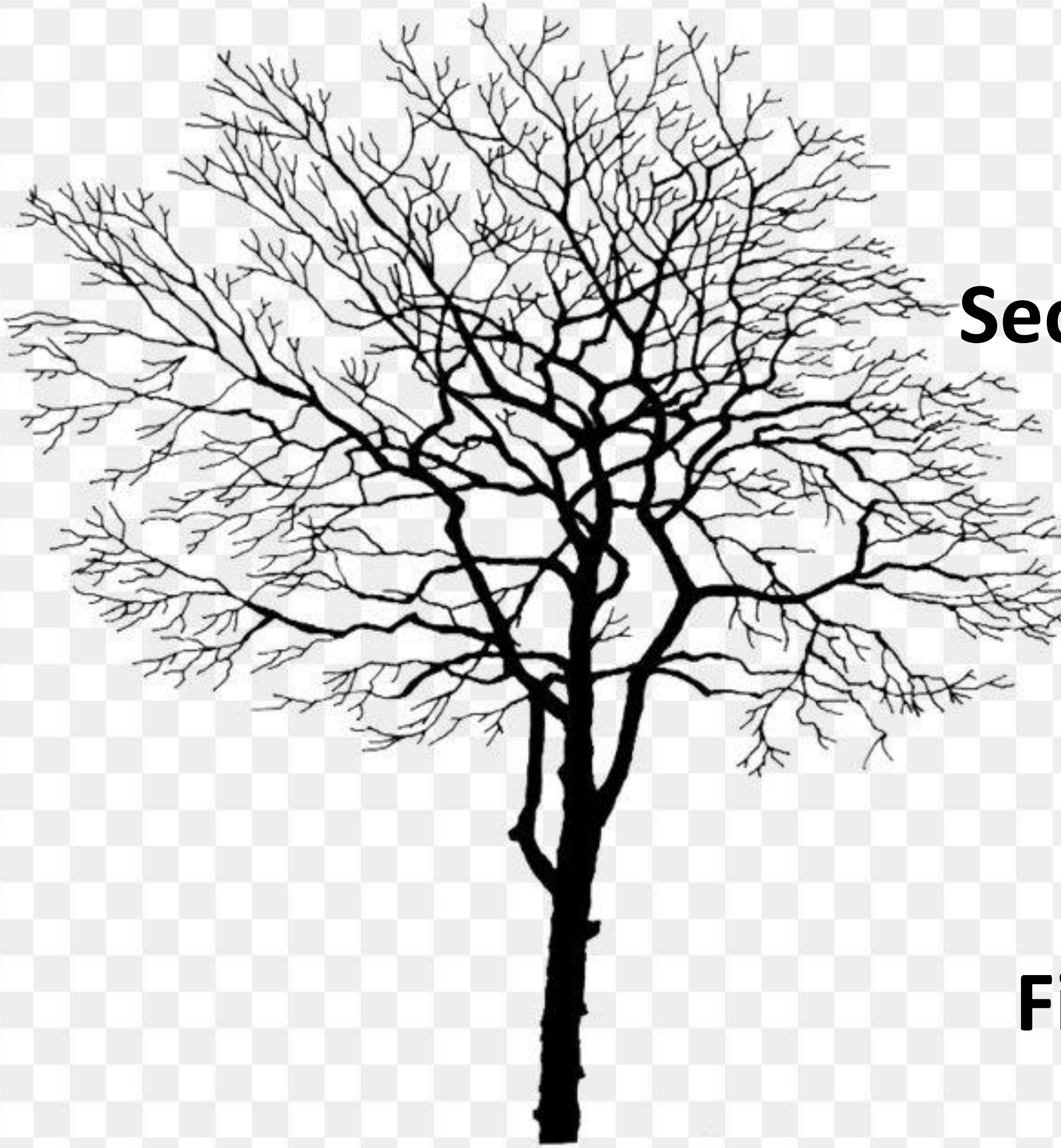
Andre Freitas

Photo by Vasilyev Alexandr

Overview of the second part of the course

Pedagogical Take

- Giles provided you with the foundations on logical representation, reasoning and programming.
- We will build and expand on it.
- Comparatively, this part of the course will cover more topics (broader strokes).
- Fundamental to provide you with an end-to-end view on Symbolic AI.
- Mastering complexity.



Second Part

First Part

How to Study

- Be in a position to define and explain:
  - The core concepts and algorithms.
  - Why they are relevant?
  - When you should use/not use them?

- Basic application of the core algorithms.

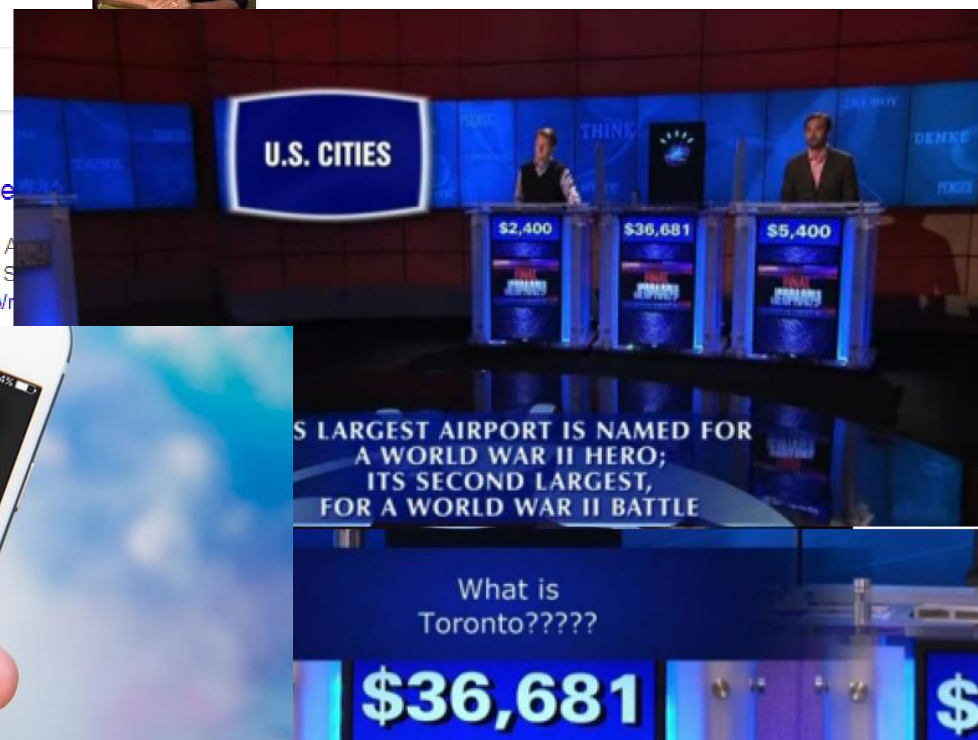
Employability Skills

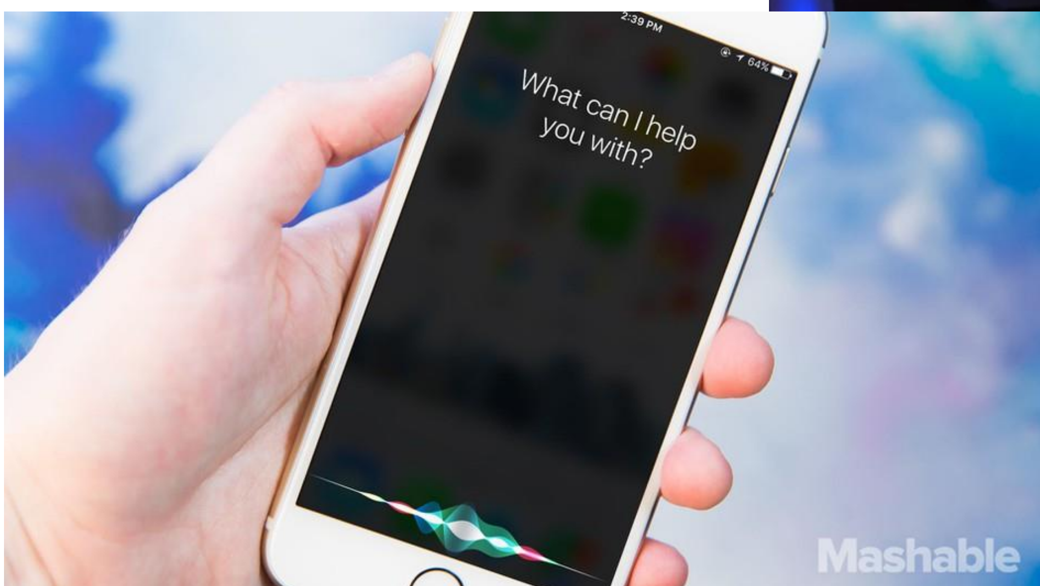
- This unit is heavily complementary to Machine Learning and Text Mining.

- Contemporary AI is evolving rapidly in the direction of hybrid <u>**neuro-symbolic models**</u>.

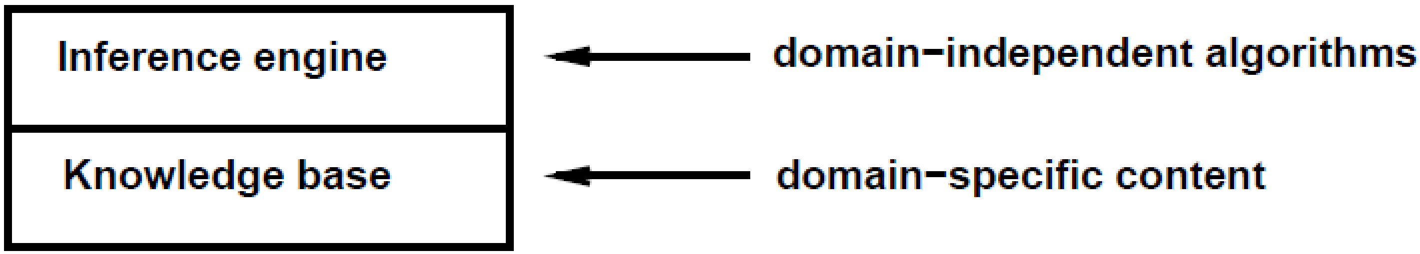
- Larger palette to build AI Systems.


Google
who is the wife of barack obama
Web
News
Images
Shopping
Videos
More
Search tools
About 186,000,000 results (0.53 seconds)
Michelle Obama (m. 1992)
Barack Obama, Spouse
Michelle Obama - Wikipedia, the free encyclope
en.wikipedia.org/wiki/Michelle_Obama
Wikipedia
Michelle LaVaughn Robinson Obama (born January 17, 1964), an A
writer, is the wife of the 44th and current President of the United S
Capers Funnye - Sidley Austin - Hyde Park, Chicago - Jeremiah Wr


U.S. CITIES
$2,400
$36,681
$5,400
S LARGEST AIRPORT IS NAMED FOR
A WORLD WAR II HERO;
ITS SECOND LARGEST,
FOR A WORLD WAR II BATTLE
What is
Toronto?????
$36,681


What can I help
you with?
Mashable

Knowledge Bases



- Knowledge base = set of sentences in a formal language
- Declarative approach to building an agent (or other system):
  - Tell it what it needs to know
- Then it can Ask itself what to do|answers should follow from the KB
- Agents can be viewed at the knowledge level
  - i.e., what they know, regardless of how implemented
- Or at the implementation level
  - i.e., data structures in KB and algorithms that manipulate them

Russell & Norvig

KB Agent

function KB-AGENT(*percept*) **returns** an *action*
  static: *KB*, a knowledge base
    t, a counter, initially 0, indicating time

  TELL(*KB*, MAKE-PERCEPT-SENTENCE(*percept*, *t*))
  action ← ASK(*KB*, MAKE-ACTION-QUERY(*t*))
  TELL(*KB*, MAKE-ACTION-SENTENCE(*action*, *t*))
  $t \leftarrow t + 1$
  return *action*

Russell & Norvig

A simple knowledge-based agent

- The agent must be able to:
  - Represent states, actions, etc.
  - Incorporate new percepts
  - Update internal representations of the world
  - Deduce hidden properties of the world
  - Deduce appropriate actions

- => sound and complete reasoning with partial information states

Russell & Norvig

Knowledge in Learning

- Task: to design agents that already know something, and are trying to learn more.

- Agents must have a learning process to gain the background knowledge in the first place
  - Learning taken place afterwards define the agent's incremental/cumulative development

- Agents can start off like normal agents
  - Gain initial knowledge through inductive learning
  - After, uses background knowledge to learn more effectively

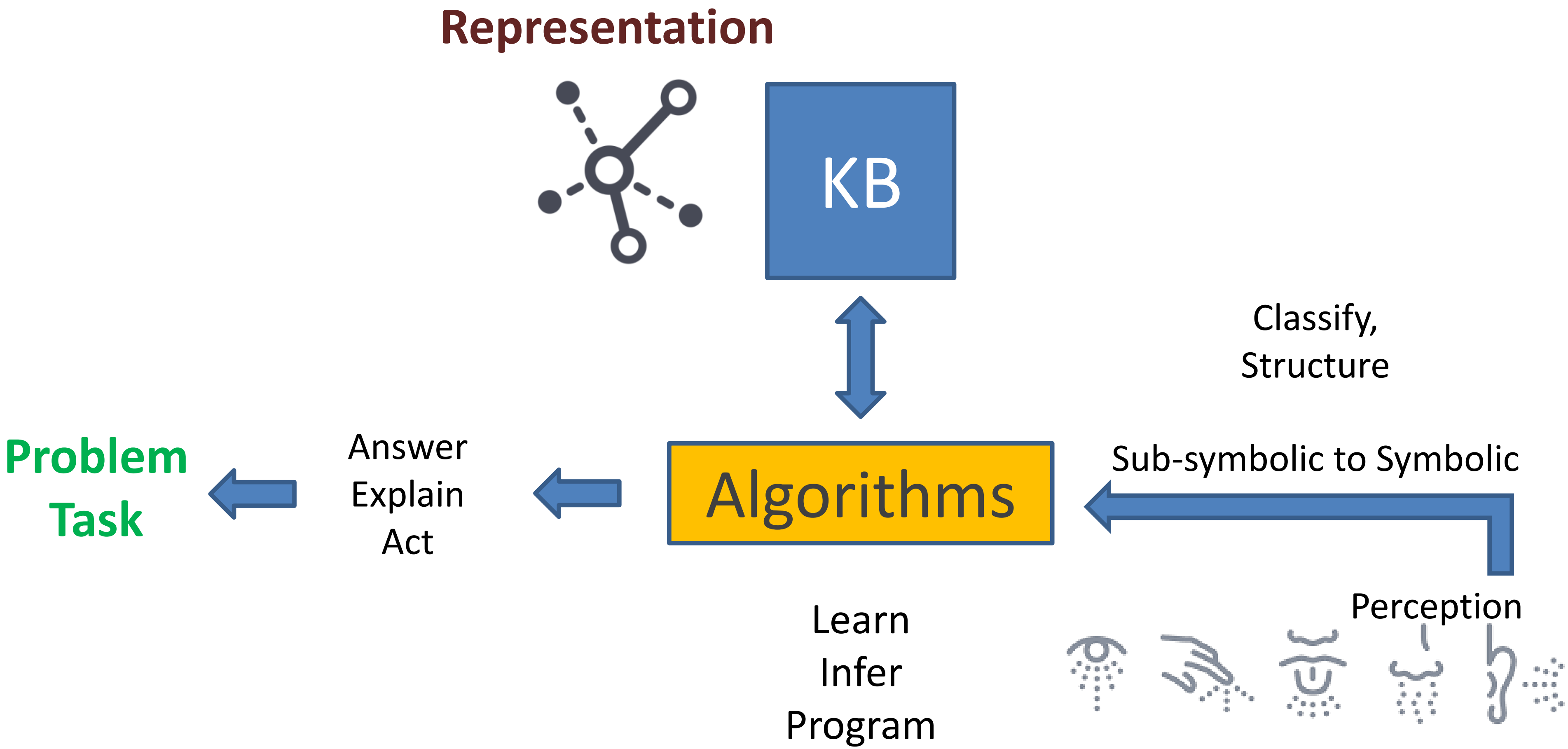
Russell & Norvig

Goals of this second part of the course

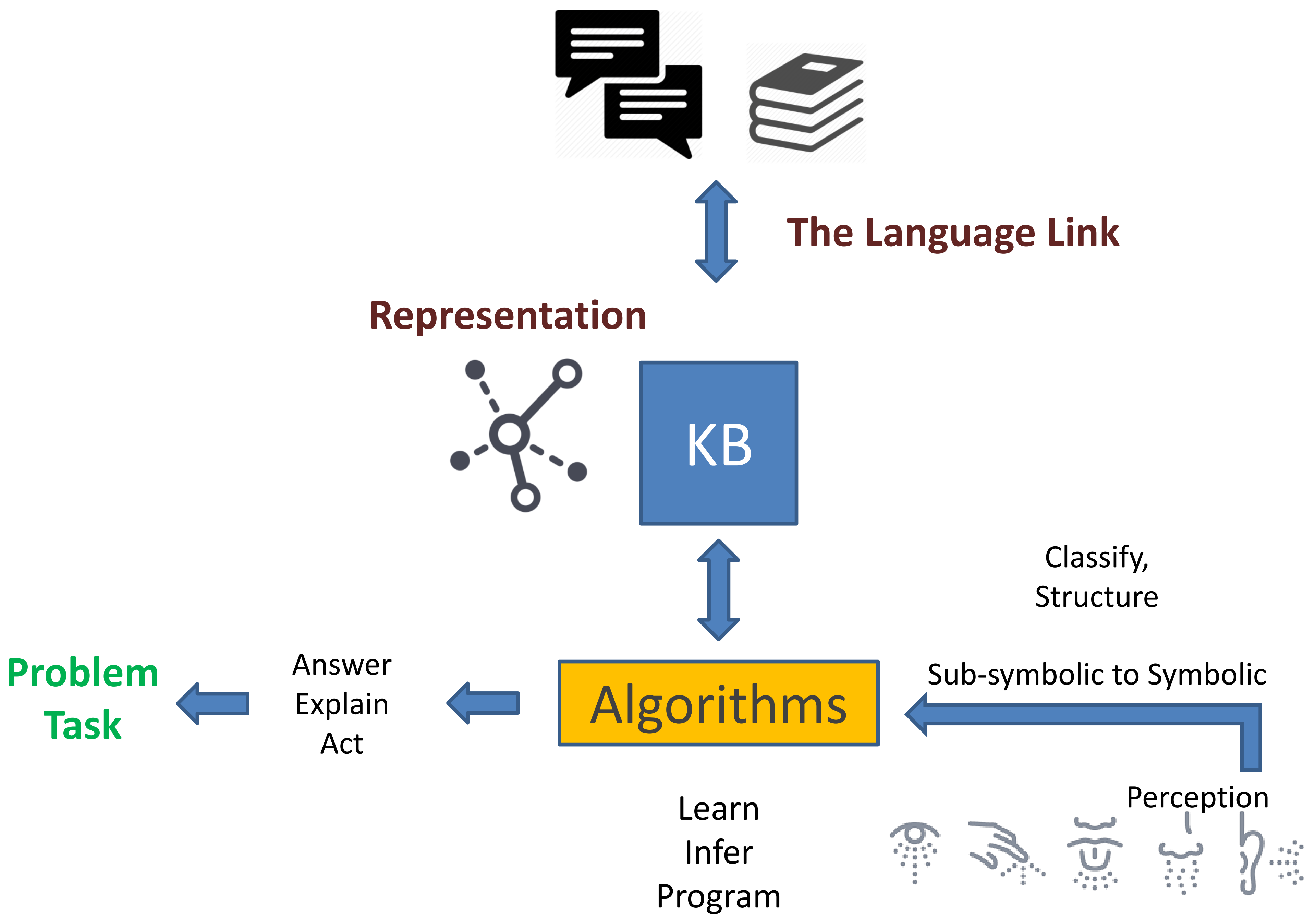
- Which knowledge representation elements support more expressive, accurate, efficient and general algorithms?

- How to build knowledge bases (under a certain representation)?

- What are the conceptual frameworks and algorithms for knowledge-based learning and inference?



Representation
KB
Classify,
Structure
Problem
Task
Answer
Explain
Act
Algorithms
Sub-symbolic to Symbolic
Perception
Learn
Infer
Program

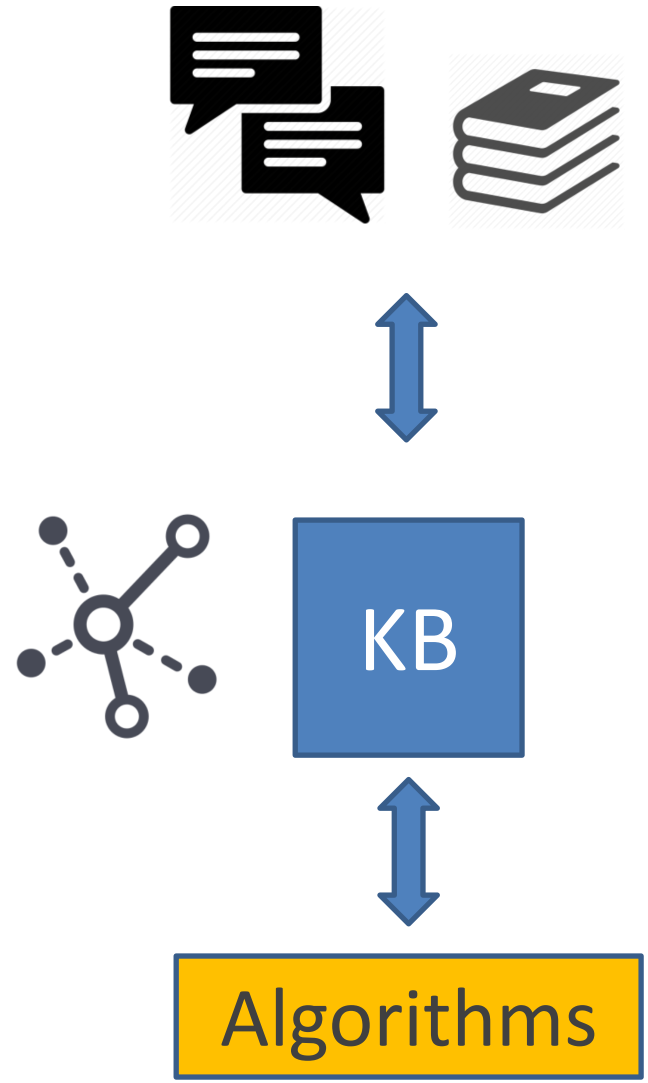

The Language Link
Representation
KB
Classify,
Structure
Problem
Task
Answer
Explain
Act
Algorithms
Sub-symbolic to Symbolic
Learn
Infer
Program
Perception

Syllabus

2. KB Construction
3. Montague Semantics
4. Semantic Parsing

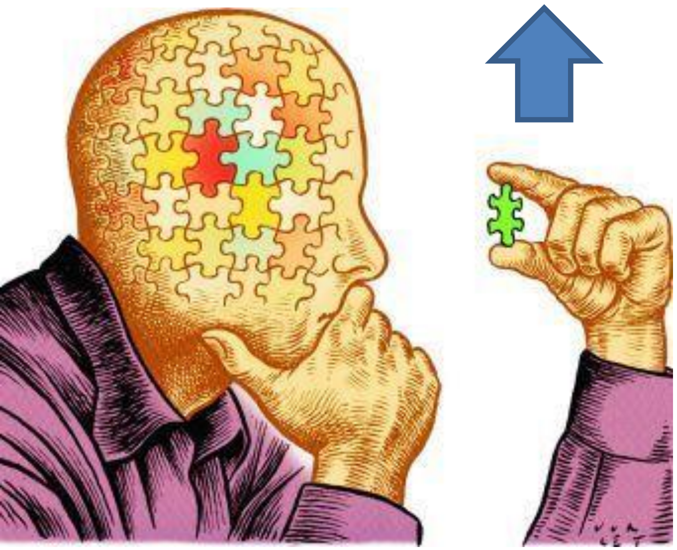
1. Knowledge Representation

5. Explanation-based Learning
6. Inductive Logic Programming
7. Natural Language Inference
8. Neuro-Symbolic Reasoning



Knowledge Representation

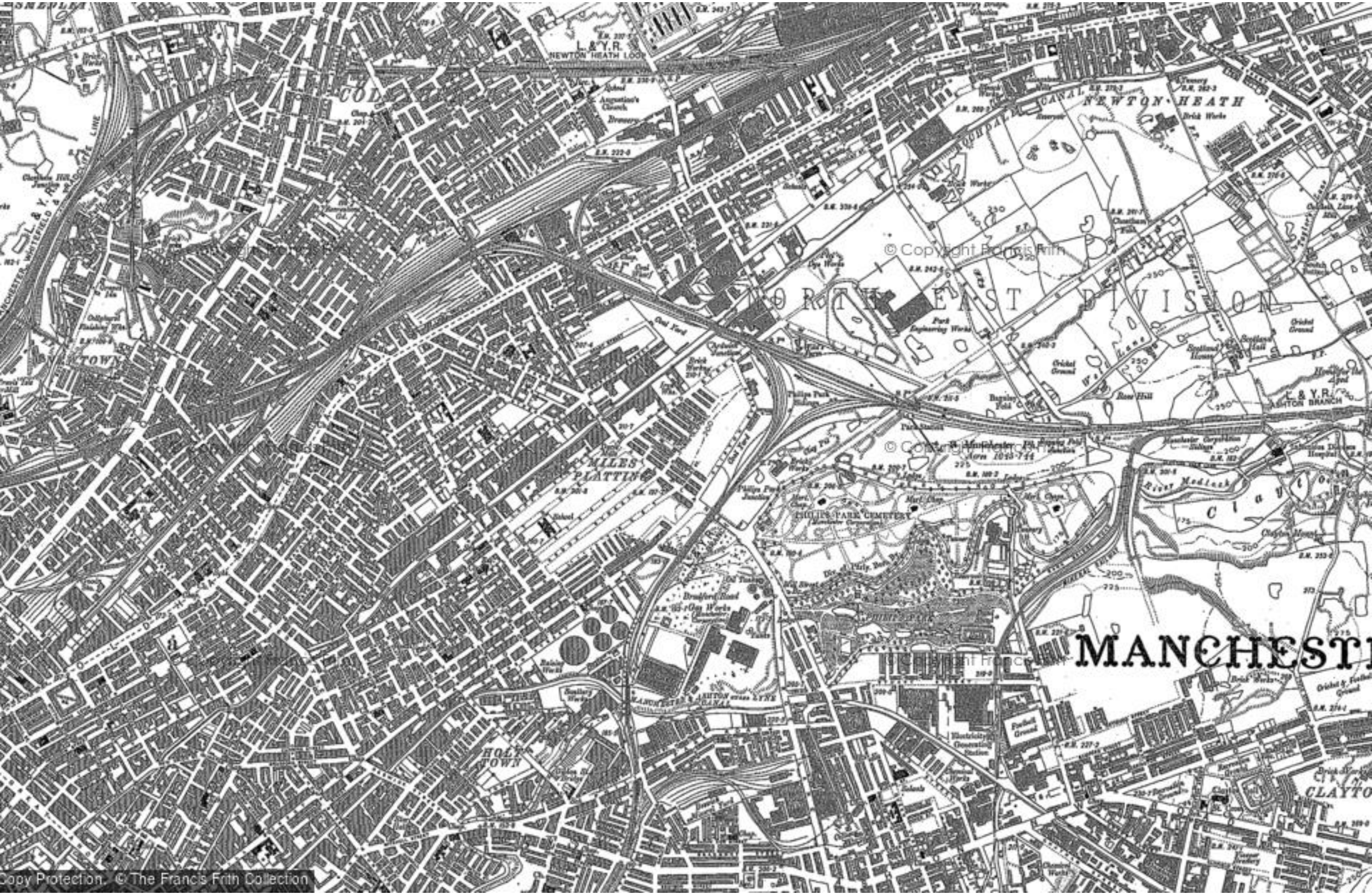
Communication of the Representation

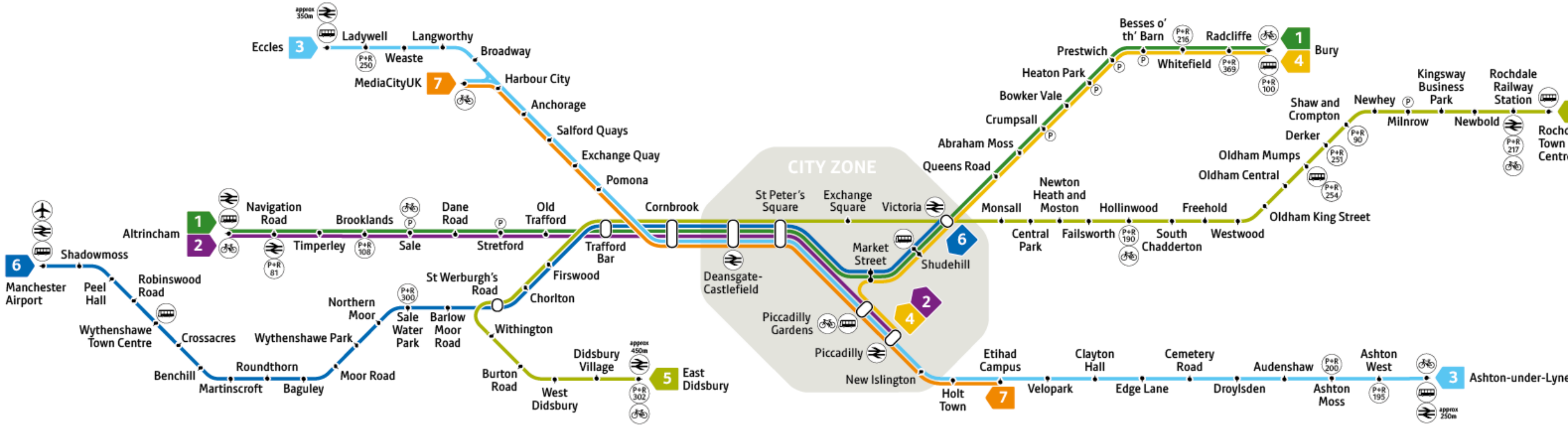


Representation of the Reality

Structure of the Reality


MANCHESTER
City Centre
Manchester Victoria
Victoria
Shudehill
Market Street
Piccadilly Gardens
St Peter's Square
Manchester Piccadilly
G-Mex
Deansgate
Manchester Oxford Road
Salford Central
Cornbrook
Wallness
New Windsor
Ancoats
Brunswick
University
Hulme
Ardwick
River Irwell
Blackfriars Road
Silk Street
Chapel Street
Deansgate
Oldfield Road
Liverpool Road
Lever Street
Oldham Road
Pollard Street
Stretford Road
Whitworth Street
Ducie Street
Store Street
Old Mill Street
Butler Street
Livesey Street
Water Street
St Stephen Street
Ordsall La
Royce Road
City Road
Brunswick Street
Every Street
Dantzic Street
Elton Street
Metrolink
A664
A6042
A6
A57
A57(M)
A5067
A34
A635
A665


L.&Y.R.
NEWTON HEATH LOCO
NEWTON HEATH
NORTH EAST DIVISION
MILES PLATTING
PHILIPS PARK CEMETERY
L.&Y.R.
ASHTON BRANCH
River Medlock
HOLT TOWN
MANCHESTE
CLAYTO



Services

1 Altrincham – Bury

2 Altrincham – Piccadilly

3 Ashton-under-Lyne – Eccles

4 Bury – Piccadilly

5 East Didsbury – Rochdale Town Centre

6 Manchester Airport – Victoria*

7 MediaCityUK – Etihad Campus

* Early services operate on a 20 minute frequency between Manchester Airport and Deansgate-Castlefield. Please check journey planning posters or **tfgm.com** before travelling.

Key

Metrolink stop

Change with other servic

Bus interchange

Rail interchange

P+R 302 Park + Ride
with number of spaces

Car park
fewer than 50 spac

Cycle Hub
membership required


Ceci n'est pas une pipe.
Magritte

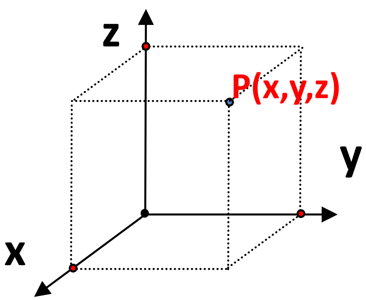
"Human knowledge is a process of approximation. In the focus of experience, there is comparative clarity. But the discrimination of this clarity leads into the penumbral background. There are always questions left over. The problem is to discriminate exactly what we know vaguely."

Alfred North Whitehead

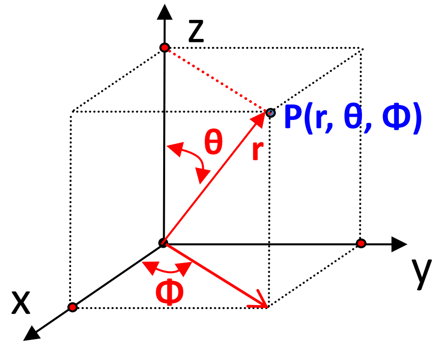
KR: Five Roles

- 1. Surrogate
  - That is, a representation
- 2. Expression of ontological commitment
  - of the world
- 3. Theory of intelligent reasoning
  - and our knowledge of it
- 4. Medium of efficient computation
  - that is accessible to programs
- 5. Medium of human expression
  - and usable

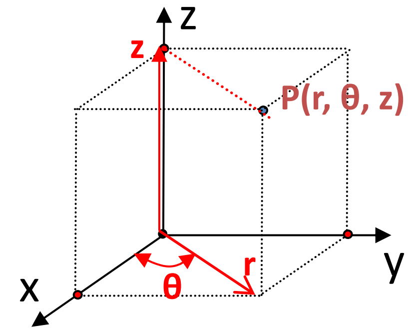
Cartesian Coordinates
P(x,y,z)



Spherical Coordinates
P(r, θ, Φ)



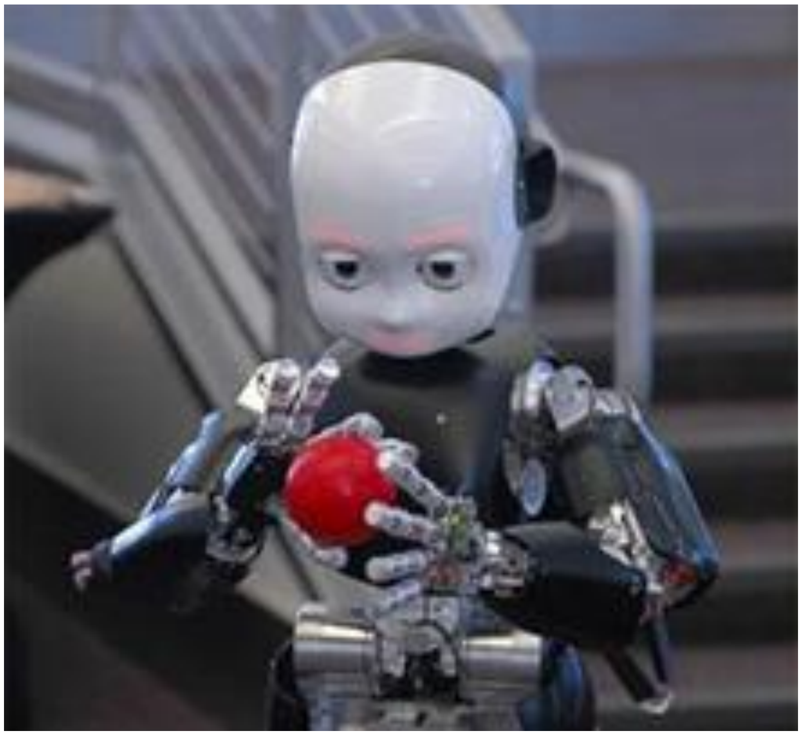
Cylindrical Coordinates
P(r, θ, z)



$$ds^2 = \left(1 - \frac{2m}{r}\right) dt^2 - \frac{1}{\left(1 - \frac{2m}{r}\right)} dr^2$$

$$-(r)^2 (d\theta^2 + \sin^2(\theta) d\phi^2)$$

Representations **deeply** impact on learning and inference



An **apple** is a sweet, edible fruit produced by an **apple tree** (***Malus pumila***).
Apple trees are cultivated worldwide and are the most widely grown species in the genus *Malus*. The tree originated in Central Asia, where its wild ancestor, *Malus sieversii*, is still found today. Apples have been grown for thousands of years in Asia and Europe and were brought to North America by European colonists. Apples have religious and mythological significance in many cultures, including Norse, Greek and European Christian traditions.

Embodied representations

Symbolic representations

"The distinctive feature of brains such as the one we own is <u>their uncanny ability to create maps</u>... But when brains make maps, they are also creating images, the main currency of our minds. Ultimately consciousness allows us to experience maps as images, to manipulate those images, and to apply reasoning to them."

Antonio Damasio (2010)

Semantics

=

Formal meaning representation model (lots of data)

+

inference model

This behaves a lot like intelligence!

> 2000 years of tradition!

Semantics

=

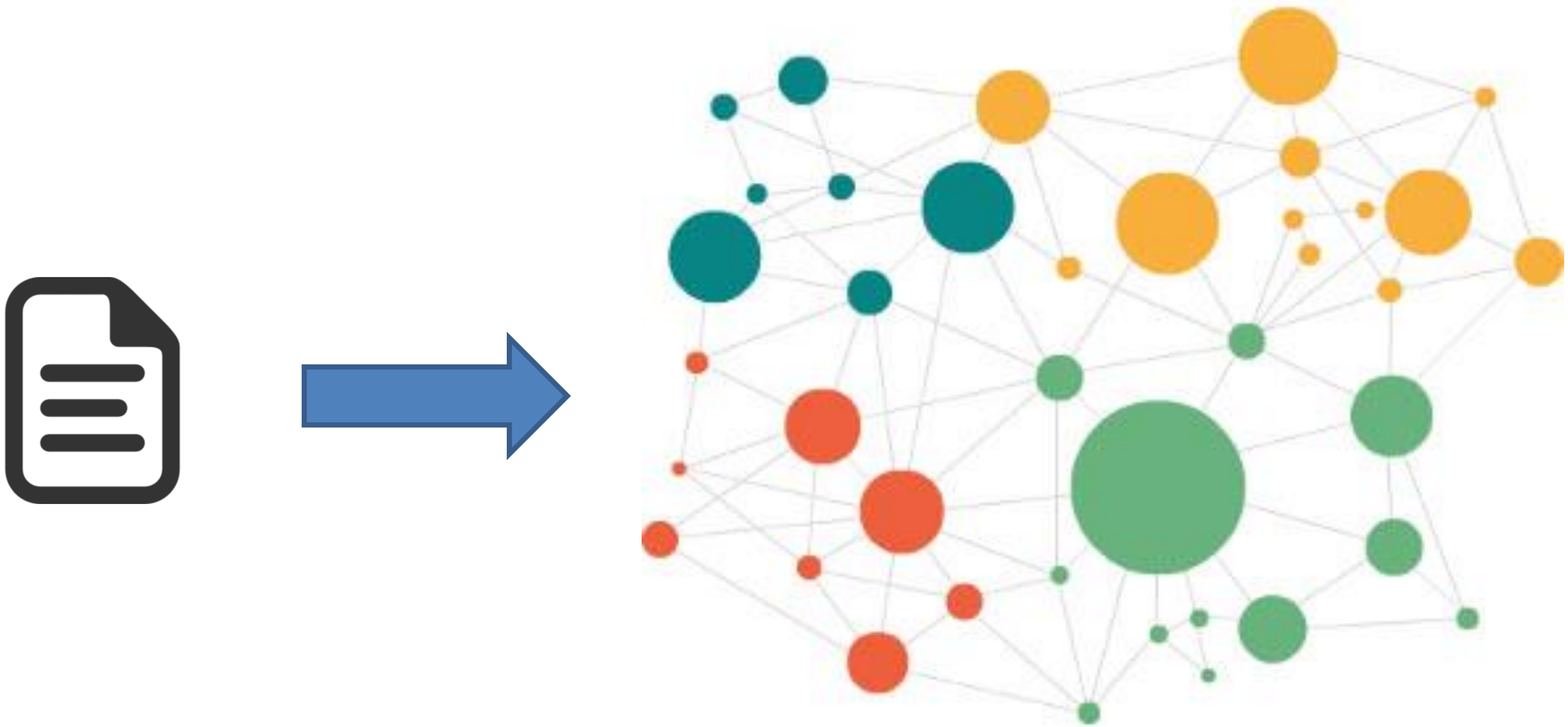
Formal meaning representation model (lots of data)

\+

inference model

Logics, linguistics, philosophy, cognitive sciences, computer science

This behaves a lot like intelligence!



Building Knowledge Bases

Data **Intelligence**

Structure/Semantics

Unstructured Data

KB Construction

Structured Data

Easy to generate

Easy to analyze (computationally)

Consistent
Comparable
Processable

From Text to Structure

...

Barack Obama went with his daughter Malia to the baseball game.

...

...

Today, during an official visit, Natasha called to her father, the president of the United States.

...

From Text to Structure

...

Barack Obama went with his daughter **Malia** to the baseball game.

...

...

Today, during an official visit, **Natasha** called to her father, the president of the United States.

...

Named Entity recognition, disambiguation

From Text to Structure

...
Barack Obama went with his daughter **Malia** to the baseball game.
...

...
Today, during an official visit, **Natasha** called to her father, the **president of the United States**.
...

Co-reference resolution

From Text to Structure

...

Barack Obama went with **his daughter** **Malia** to the baseball game.

...

...

Today, during an official visit, **Natasha** called to **her father**, the **president of the United States**.

...

Relation extraction

Regularities in Natural Language

Barack Obama went with his daughter Malia to the baseball game.

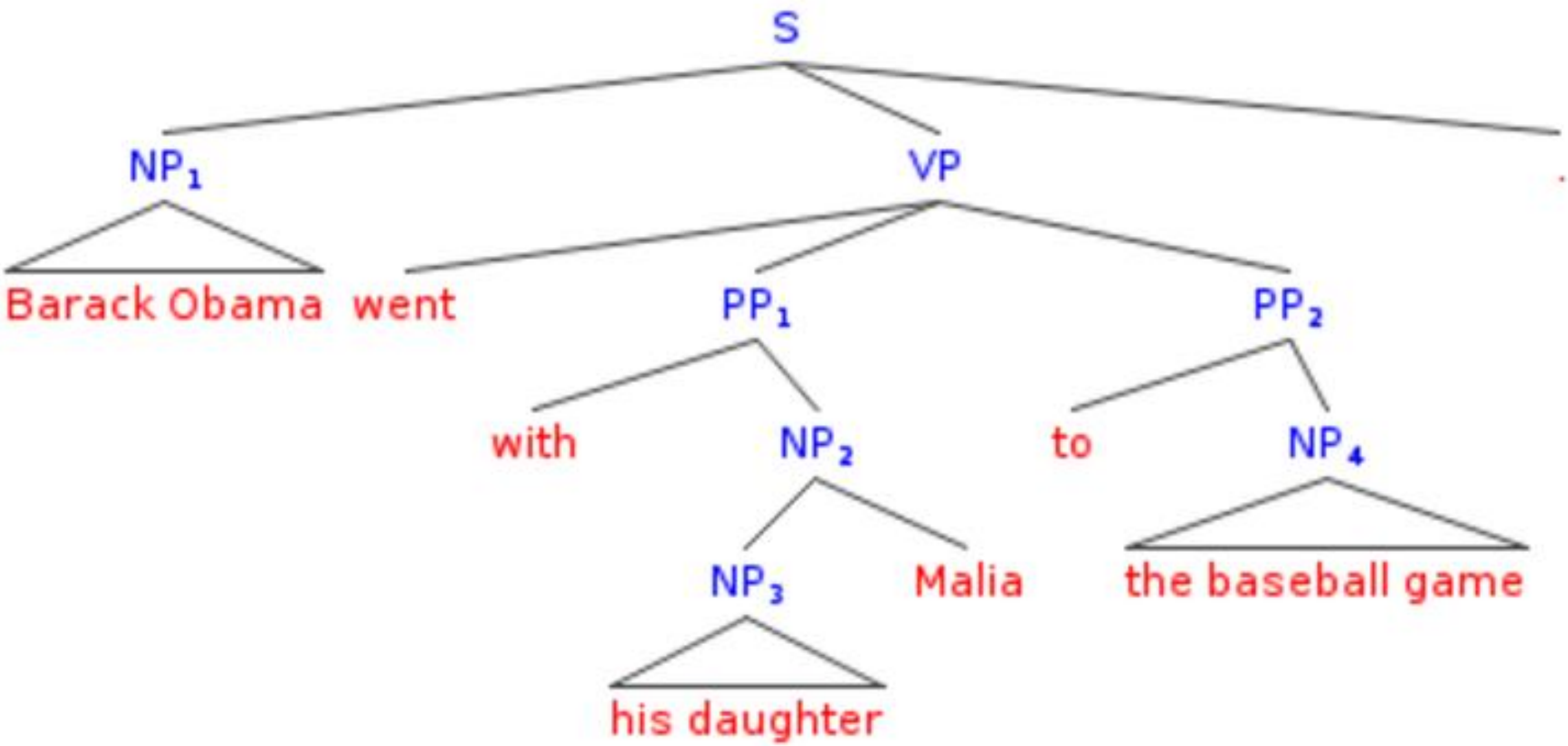
Barack/NNP
Obama/NNP
went/VBD
with/IN
his/PRP$
daughter/NN
Malia/NN
to/TO
the/DT
baseball/NN
game/NN
./.

Regularities in Natural Language

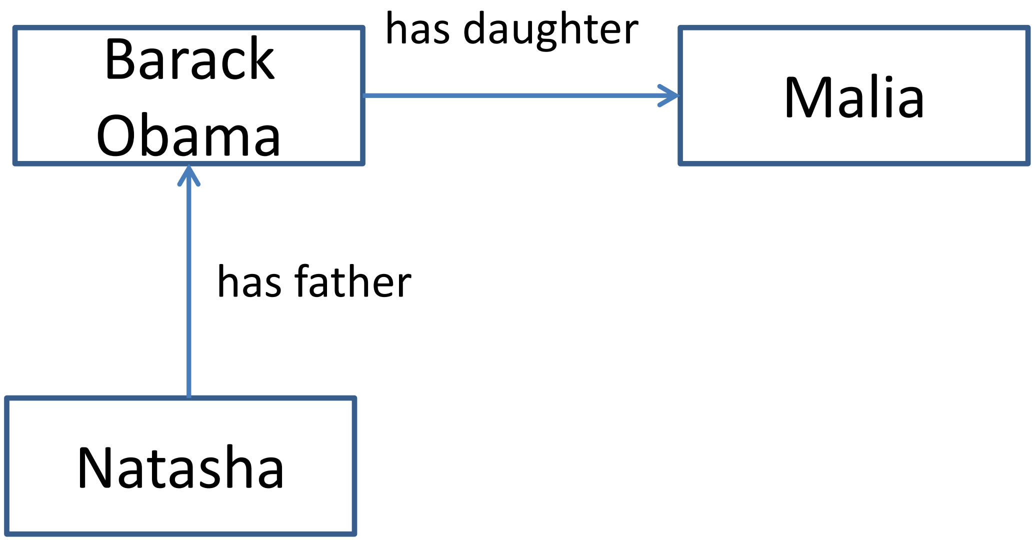
Barack Obama went with his daughter Malia to the baseball game.

Barack/**NNP** → Proper noun
Obama/**NNP**
went/VBD
with/IN
his/PRP$
daughter/NN
Malia/NN
to/TO
the/DT
baseball/NN
game/NN
./.

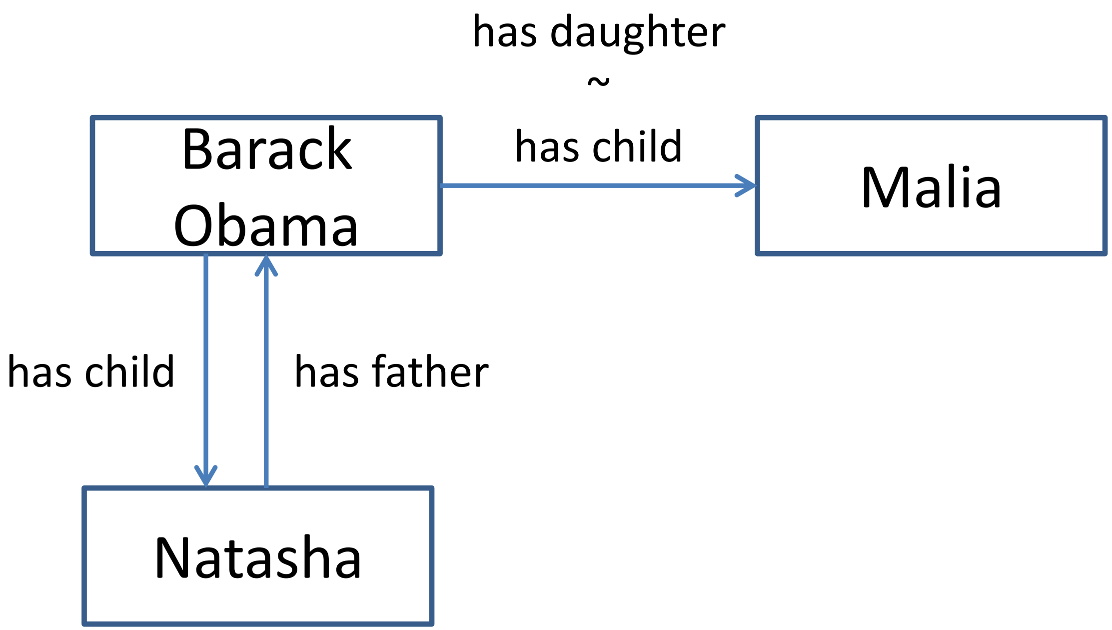
Regularities in Natural Language



Structural/Logical Form



Structural/Logical Form



Applying some logical or corpus-based inference we get ....

Rephrasing it

- has_child(Barack_Obama, Malia)
- has_child(Barack_Obama, Natasha)
- …

Now we can answer this query

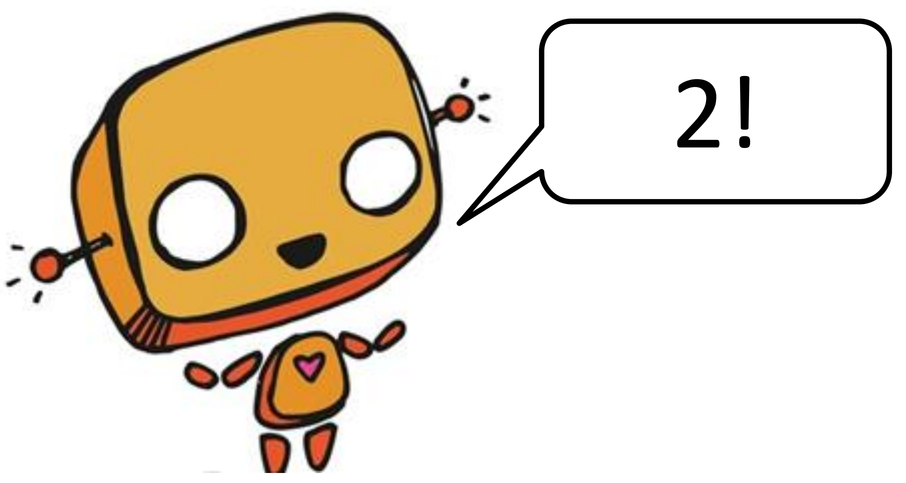
- *How many children does Barack Obama have?*

some magic called **semantic parsing** goes on ...

- count(has_child(Barack_Obama, ?x))

It computes!

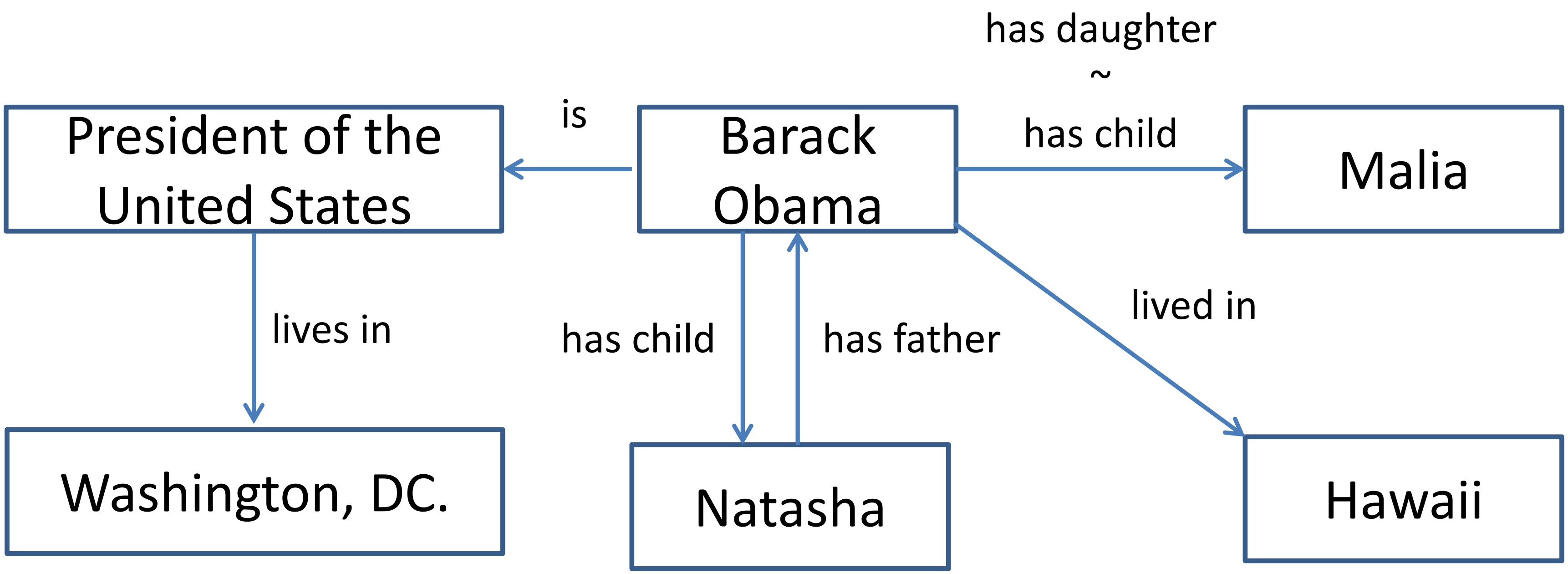
Query: count(has_child(Barack_Obama, ?x))



KB:

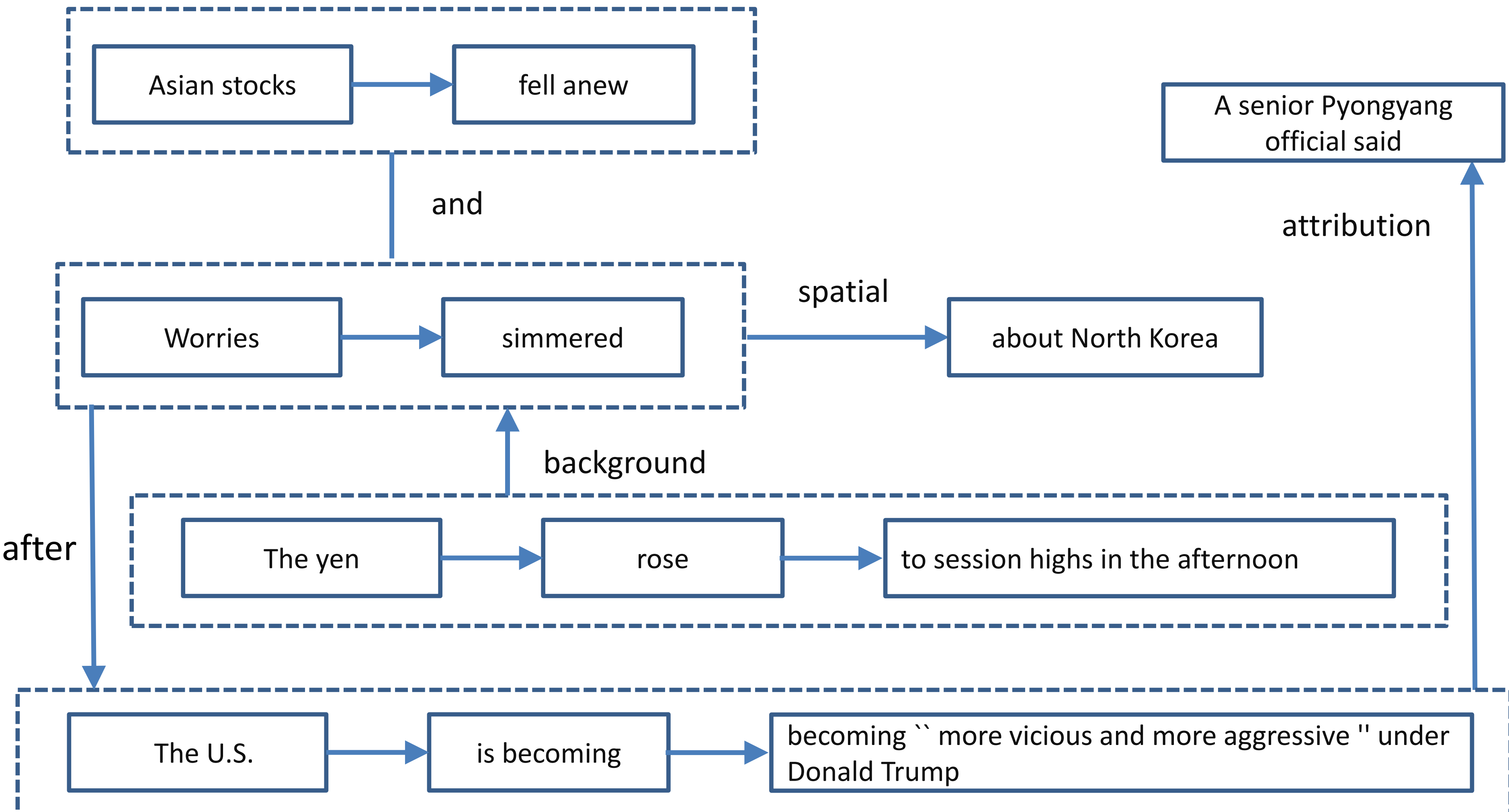
- has_child(Barack_Obama, Malia)
- has_child(Barack_Obama, Natasha)
- …

Extrapolating



Semantic representation

Asian stocks fell anew and the yen rose to session highs in the afternoon as worries about North Korea simmered, after a senior Pyongyang official said the U.S. is becoming ``more vicious and more aggressive'' under President Donald Trump .



Semantic Parsing using CCGs

$$
\begin{array}{cccc}
\text{show} \quad \text{me} & \text{flights} & \text{to} & \text{Boston} \\
\hline
S/N & N & PP/NP & NP \\
\lambda f.f & \lambda x.flight(x) & \lambda y.\lambda x.to(x,y) & BOSTON
\end{array}
$$

$$
\xrightarrow{\hspace{6em}} >
$$

$$
\begin{array}{c}
PP \\
\lambda x.to(x, BOSTON) \\
\hline
N \backslash N \\
\lambda f.\lambda x.f(x) \wedge to(x, BOSTON)
\end{array}
$$

$$
\xrightarrow{\hspace{6em}} <
$$

$$
\begin{array}{c}
N \\
\lambda x.flight(x) \wedge to(x, BOSTON)
\end{array}
$$

$$
\xrightarrow{\hspace{6em}} >
$$

$$
\begin{array}{c}
S \\
\lambda x.flight(x) \wedge to(x, BOSTON)
\end{array}
$$

Artzi

Symbol

Tree

Tree is the name of a set

Tree is a noun

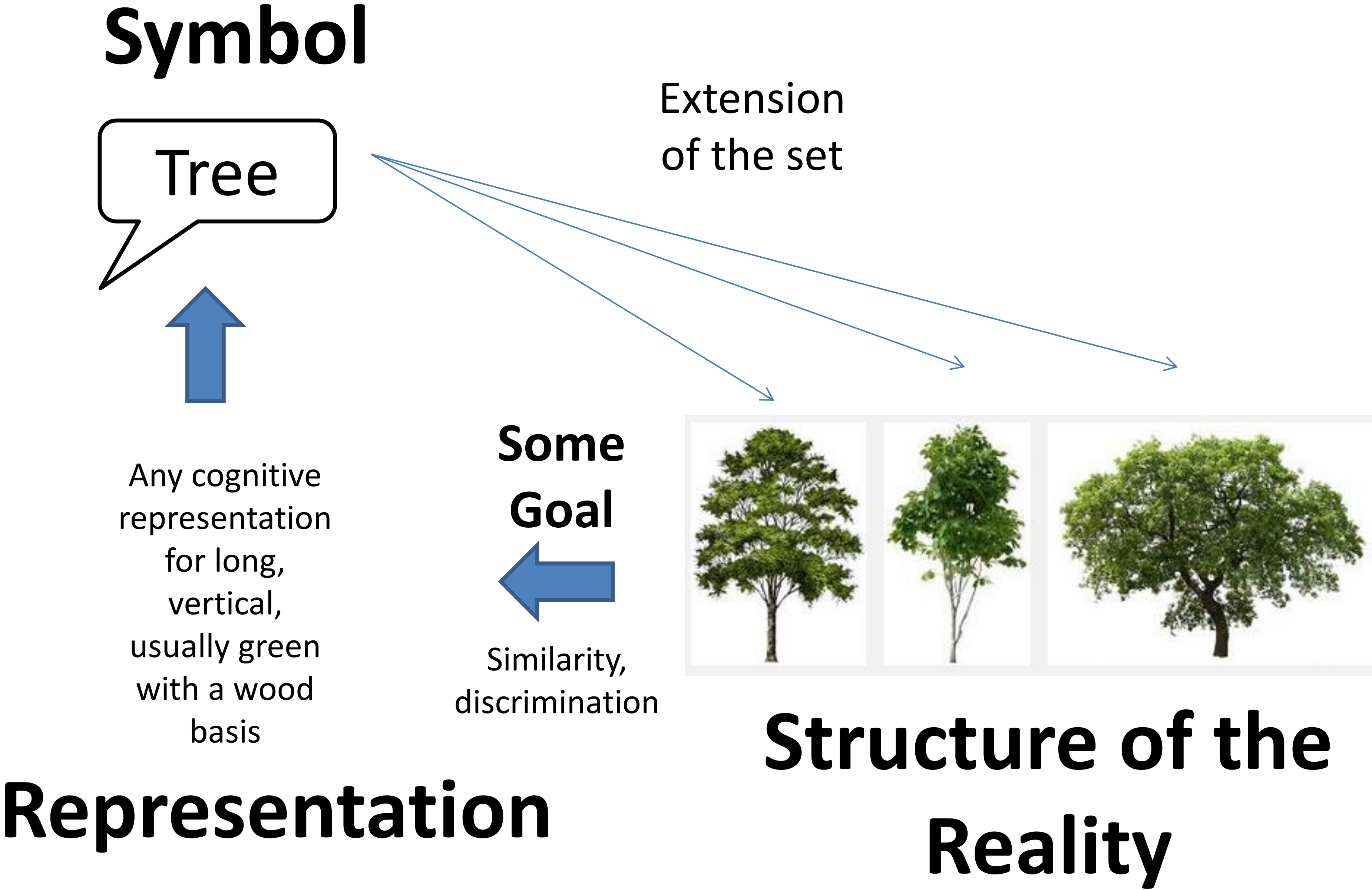
Any cognitive representation for long, vertical, usually green with a wood basis

Some Goal

Similarity, discrimination

Representation

Structure of the Reality


Symbol
Tree
Extension
of the set
Some
Goal
Any cognitive
representation
for long,
vertical,
usually green
with a wood
basis
Similarity,
discrimination
Representation
Structure of the
Reality

Operating on the Representation

Symbol

Tree

The tallest tree

Any cognitive representation for long, vertical, usually green with a wood basis

Some Goal

Similarity, discrimination

Structure of the Reality

Representation

Operating on the Representation

Definite article: "get me one"

Symbol

Tree

The tallest tree

Any cognitive representation for long, vertical, usually green with a wood basis

Some Goal

Similarity, discrimination

Structure of the Reality

Representation

Operating on the Representation

Superlative adjective : "top most"

Symbol

Tree

The **tallest** tree

Some Goal

Any cognitive representation for long, vertical, usually green with a wood basis

Similarity, discrimination

Representation

Structure of the Reality

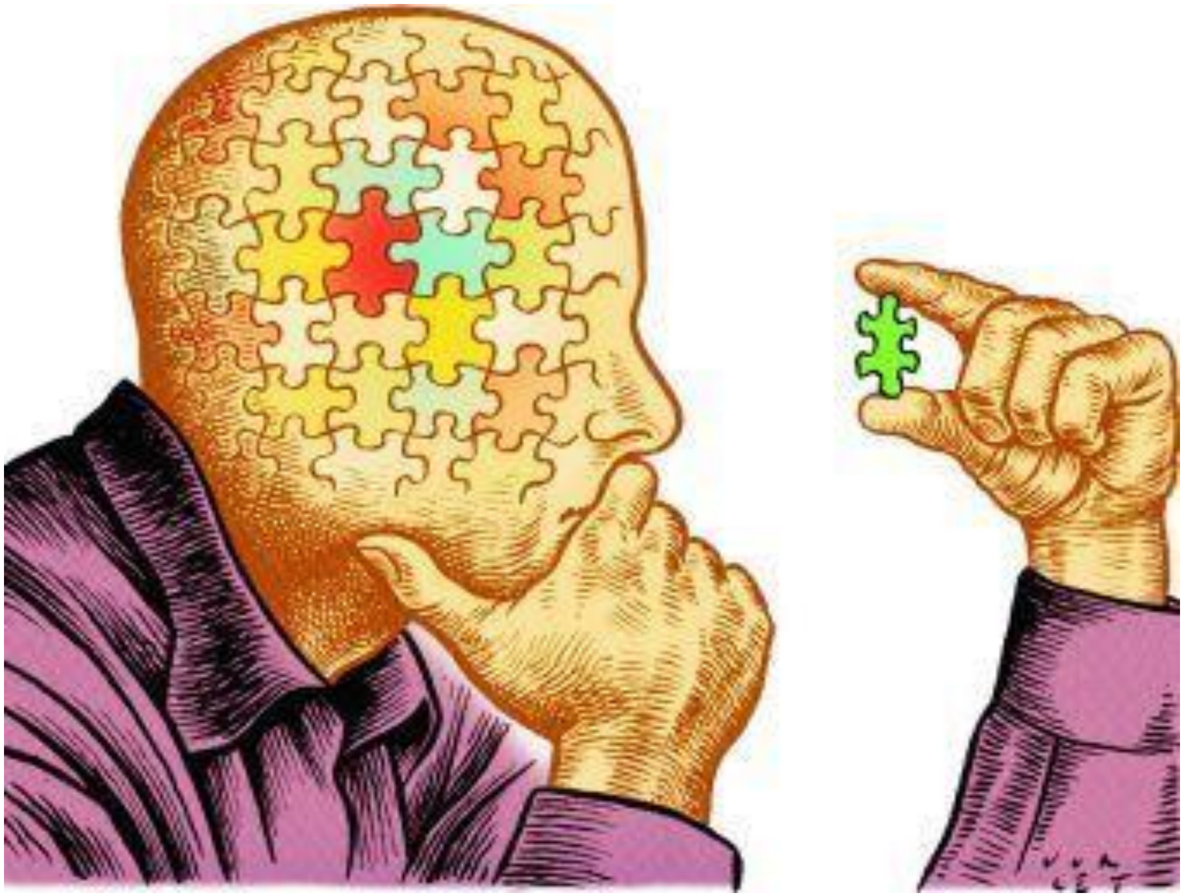
Natural Language ⟺ Representation

- There is a mapping between natural language and knowledge representation.

- Looking at natural language is looking at the representation (constrained by the communication medium).

Language as a ‘Geological outcrop’ of our cognitive representation.

https://www.geocaching.com/geocache/GC3YMK9_one-day-geology-of-oman-1?guid=c3516272-9eca-4c10-ae14-8dabd9346b98



Knowledge in
Learning and Inference

Inductive Logic Programming

- ILP algorithms are constructive induction algorithms
  - Able to create new predicates to facilitate the expression of explanatory hypotheses

- Express Grandparent
  - Empty background
  - Hypotheses are long and complicated

$$
\begin{aligned}
\text{Grandparent}(x, y) \Leftrightarrow \; & [\exists z\ \text{Mother}(x, z) \wedge \text{Mother}(z, y)] \\
\vee\; & [\exists z\ \text{Mother}(x, z) \wedge \text{Father}(z, y)] \\
\vee\; & [\exists z\ \text{Father}(x, z) \wedge \text{Mother}(z, y)] \\
\vee\; & [\exists z\ \text{Father}(x, z) \wedge \text{Father}(z, y)]
\end{aligned}
$$

Chu

Inductive Logic Programming

- By creating a new predicate, the definition of Grandparent can be reduced

$$Parent(x, y) \Leftrightarrow [Mother(x, y) \vee Father(x, y)]$$
$$Grandparent(x, y) \Leftrightarrow [\exists z\ Parent(x, z) \wedge Parent(z, y)]$$

- Background knowledge can reduce the size of hypotheses required to explain the observations

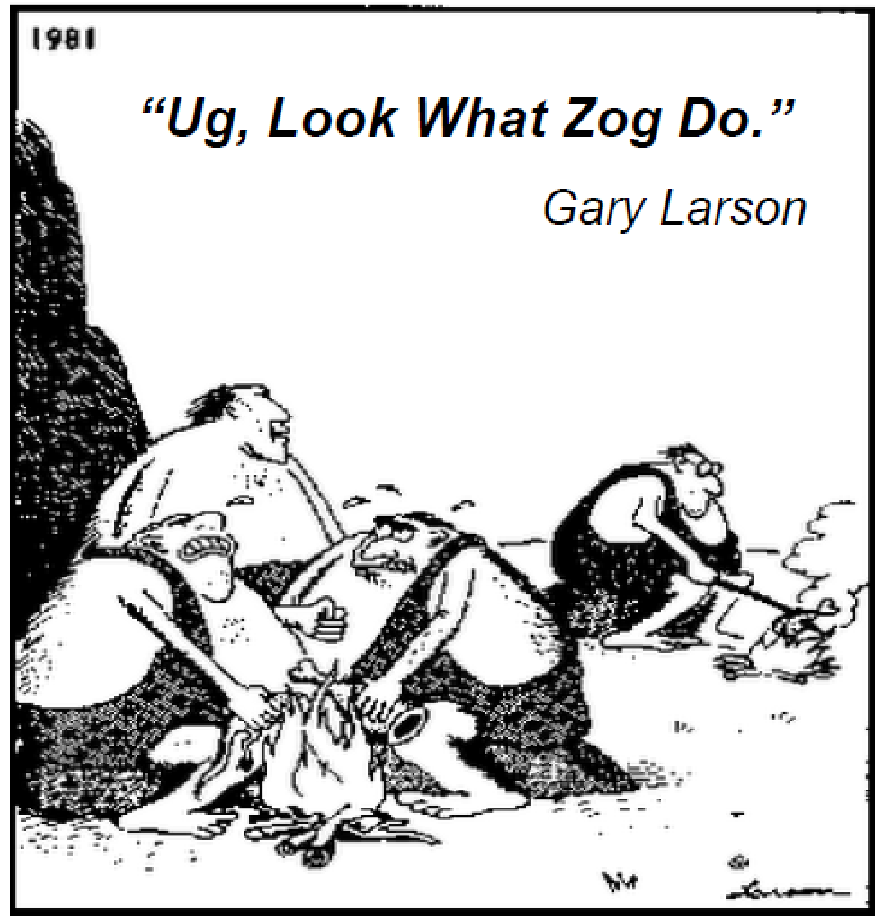
Explanation-based Learning

- **Explanation-based Learning (EBL)**
  – Method for extracting rules from individual observations through an explanation.

- **Explanation**
  – Stick holds the food over the fire while keeping hands safe.

- **Generalization**
  – Any long, rigid, sharp object can be used to toast food over the fire.

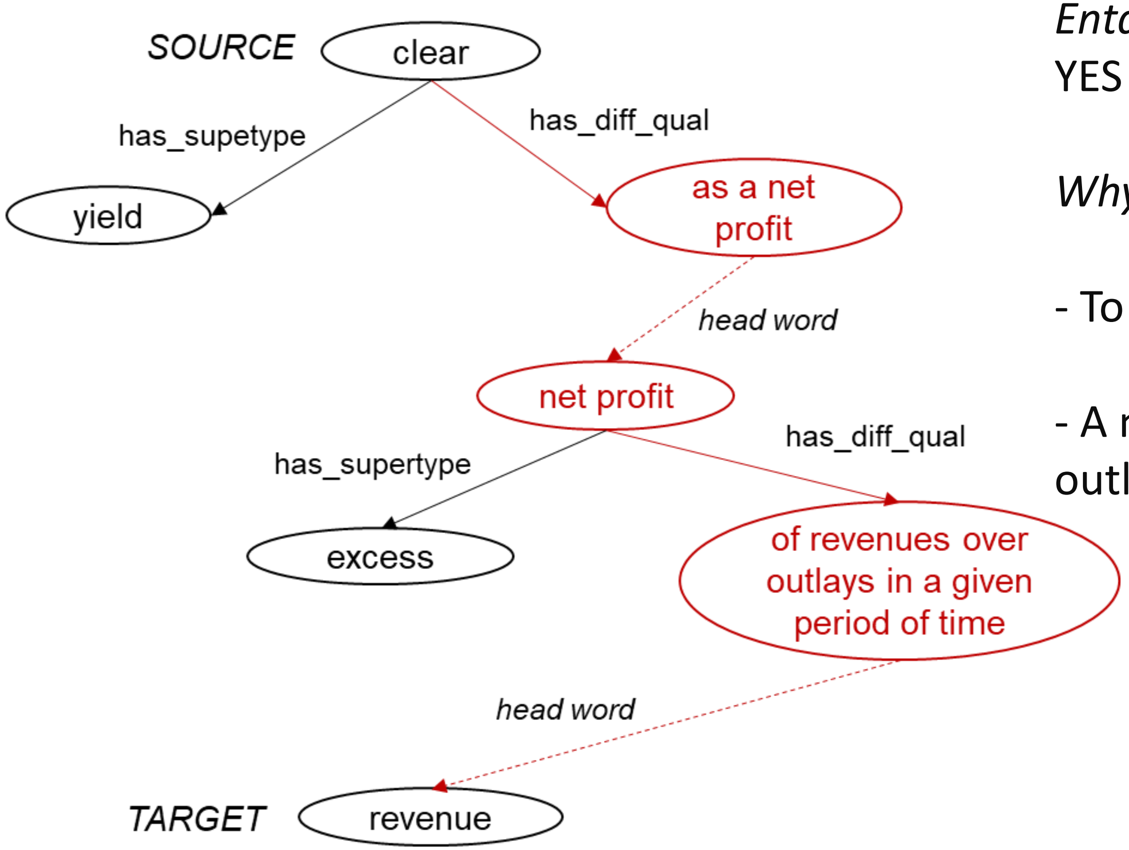
  – General rule follows logically from the background knowledge of the cavemen's usual cooking process.



Chu

Natural Language Inference

T: IBM cleared $18.2 billion in the first quarter.
H: IBM's revenue in the first quarter was $18.2 billion.



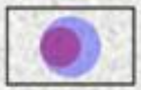
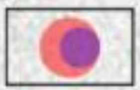
Entailment?
YES

Why?

- To clear is to yield as a net profit

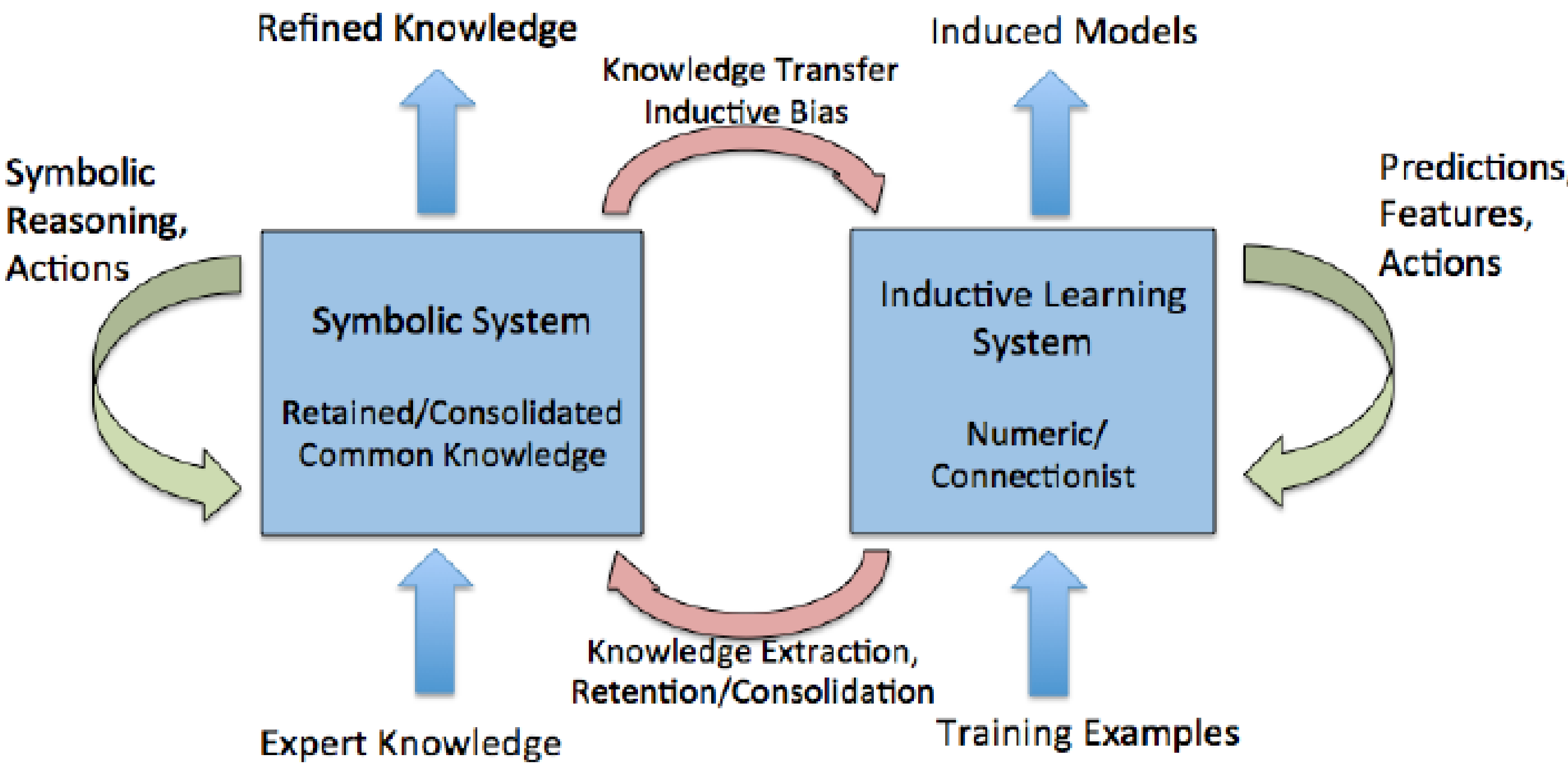
- A net profit is an excess of revenues over outlays in a given period of time

Natural Language Inference

diagram	symbo	name	example
	$x \equiv y$	equivalence	*couch* ≡ *sofa*
	$x \sqsubset y$	forward entailment (strict)	*crow* ⊏ *bird*
	$x \sqsupset y$	reverse entailment (strict)	*European* ⊐ *French*
	x ^ y	negation (exhaustive exclusion)	*human* ^ *nonhuman*
	$x \mid y$	alternation (non-exhaustive exclusion)	*cat* | *dog*
	x □ y	cover (exhaustive non-exclusion)	*animal* □ *nonhuman*
	x # y	independence	*hungry* # *hippo*

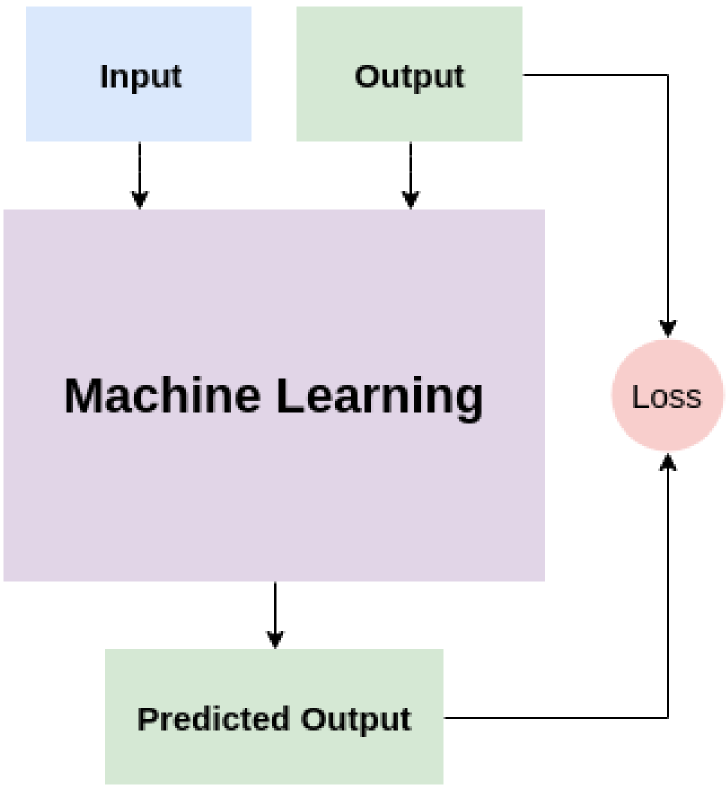
MacCartney

Neuro-Symbolic Models



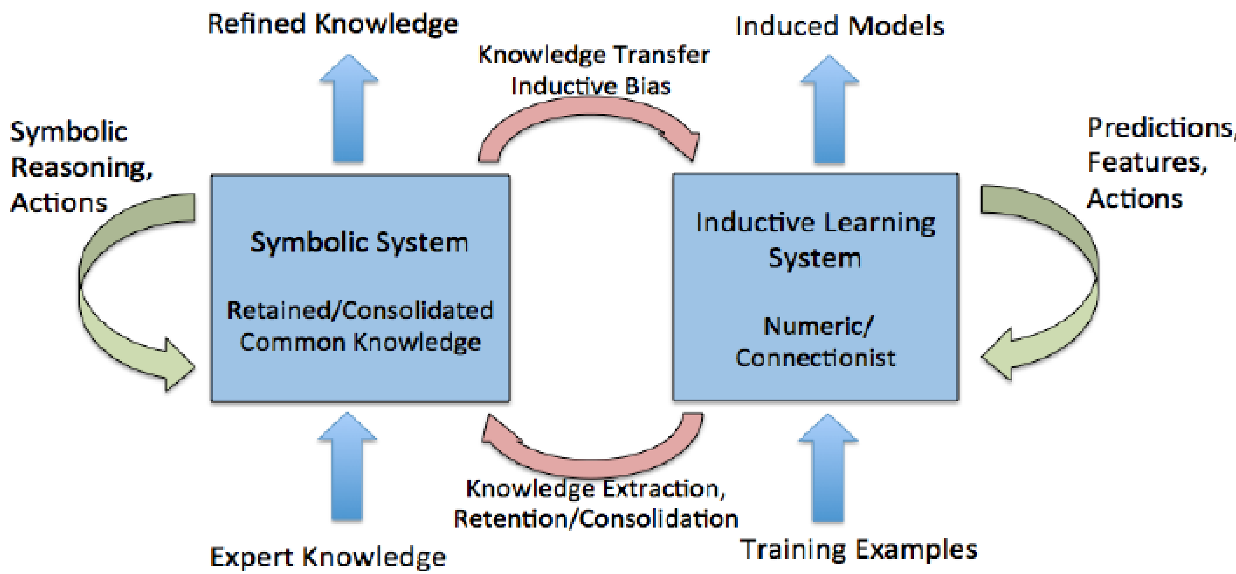
Current Limits of Deep Learning

1. Deep learning thus far is data hungry
2. Deep learning thus far is shallow and has limited capacity for transfer
3. Deep learning thus far has no natural way to deal with hierarchical structure
4. Deep learning thus far is not sufficiently transparent
5. Deep learning thus far has not been well integrated with prior knowledge
6. Deep learning thus far cannot inherently distinguish causation from correlation
7. Deep learning presumes a largely stable world, in ways that may be problematic
8. Deep learning thus far works well as an approximation, but its answers often cannot be fully trusted
9. Deep learning thus far is difficult to engineer with



Marcus, Gary. "Deep learning: A critical appraisal." *arXiv preprint arXiv:1801.00631* (2018).

Statistical vs Symbolic AI Systems



	Statistical	Symbolic
Explainability	Hard	Easy
Generalizing algebraic operations	Hard	Easy
Robustness to noise	Easy	Hard
Robustness to ambiguity	Easy	Hard
Robustness to mislabeling	Easy	Hard

∂ILP

It is possible for systems to combine statistical perceptual with conceptual interpretable reasoning!

	Deep Learning	Symbolic Program Synthesis	∂ILP
Robust to noise	YES	NO	YES
Can learn from non-symbolic data	YES	NO	YES
Data efficient	NO	YES	YES
Interpretable	NO	YES	YES

Evans, Richard, and Edward Grefenstette. "Learning explanatory rules from noisy data." *Journal of Artificial Intelligence Research*61 (2018): 1-64.

Gated Graph Neural Networks

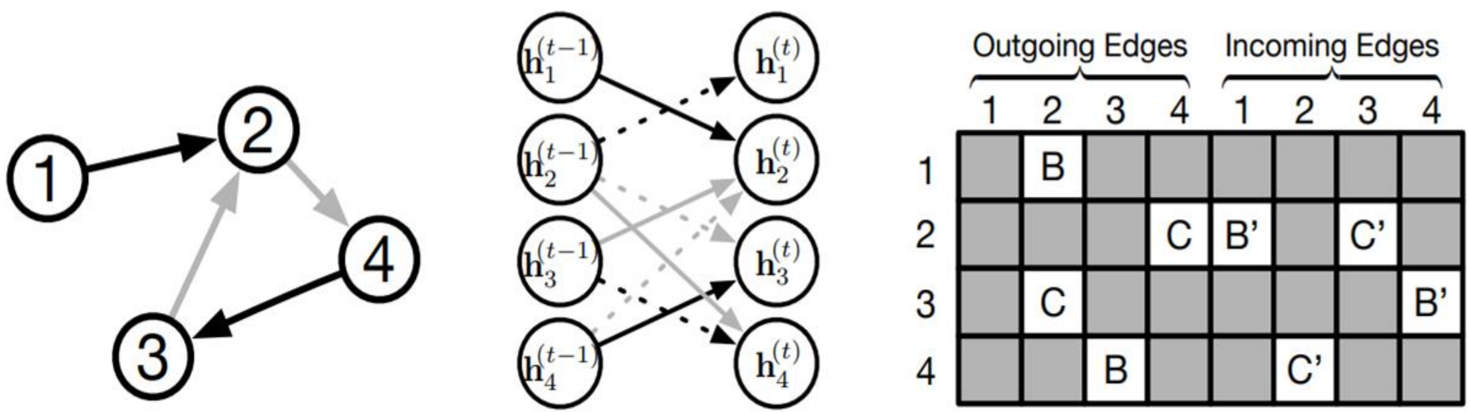
GGNNs are a neural network architecture defined according to a graph structure G = (V, E)

Nodes v ∈ V take unique values from 1, . . . , |V|, and edges are pairs e = (v, v0) ∈ V × V

GGNNs map graphs to outputs via two steps. First, there is a **propagation step** that computes node representations for each node; second, an **output model** ov = g(hv, lv) maps from node representations and corresponding labels to an output ov for each v ∈ V

The propagation model is similar to an LSTM. Each node in the graph v has a hidden state representation h(t)v that is updated at every time step t. The computation starts at t = 0 with initial hidden states xv that depends on the problem.

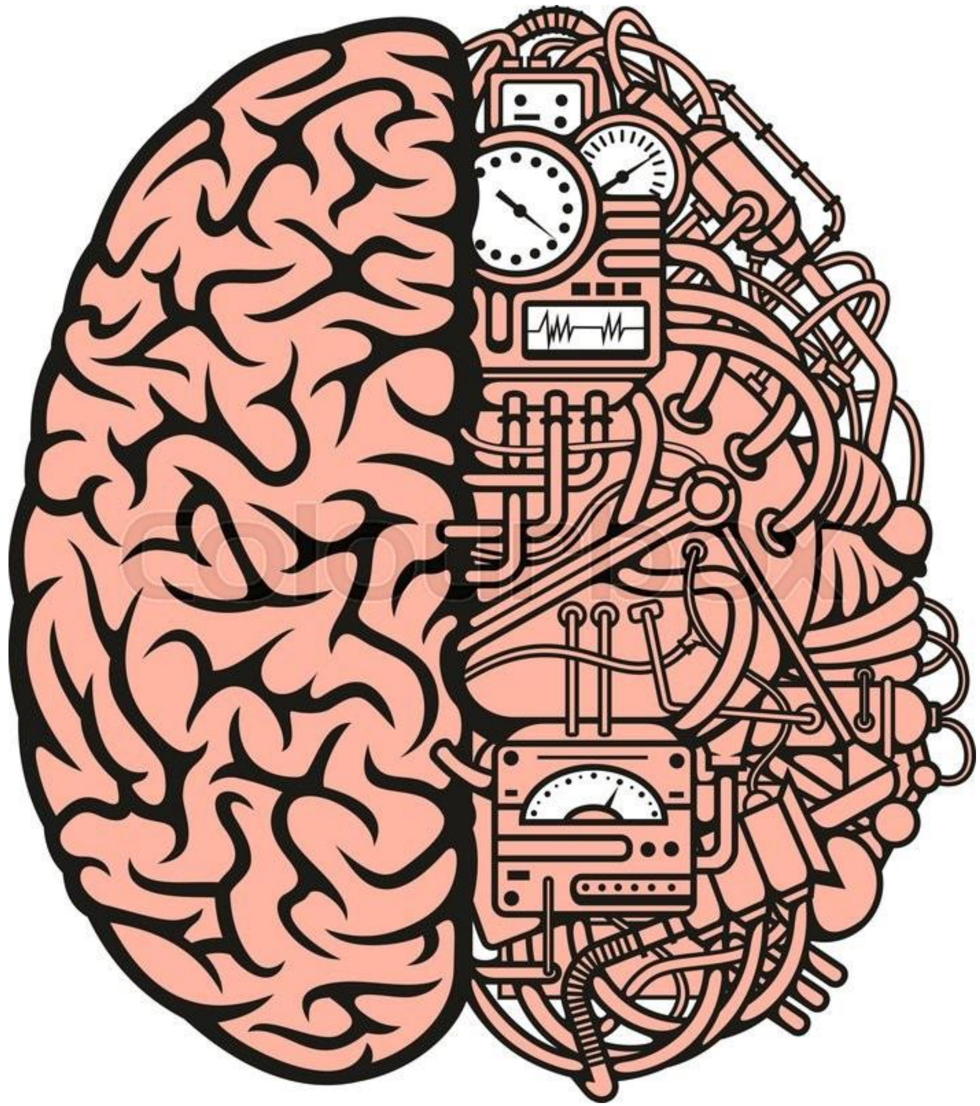
The structure of the graph, encoded in a matrix A serves to retrieve the hidden states of adjacent nodes based on the edge types between them. The hidden states are then updated by a a gated update module.



(a) (b) (c) $\mathbf{A} = \left[\mathbf{A}^{(\mathrm{out})}, \mathbf{A}^{(\mathrm{in})}\right]$

Li, Yujia, et al. "Gated graph sequence neural networks." *arXiv preprint arXiv:1511.05493* (2015).

Building Neuro-Symbolic Systems



Explainable QA

Q: The Pollution Prevention Act of 1990 expanded a publicly available or private database?

A: publicly available

Explanation:

Paragraph A: Pollution Prevention Act of 1990

The Pollution Prevention Act of 1990 (PPA) in the United States created a national policy to have pollution prevented or reduced at the source wherever possible. It also expanded the Toxics Release Inventory.

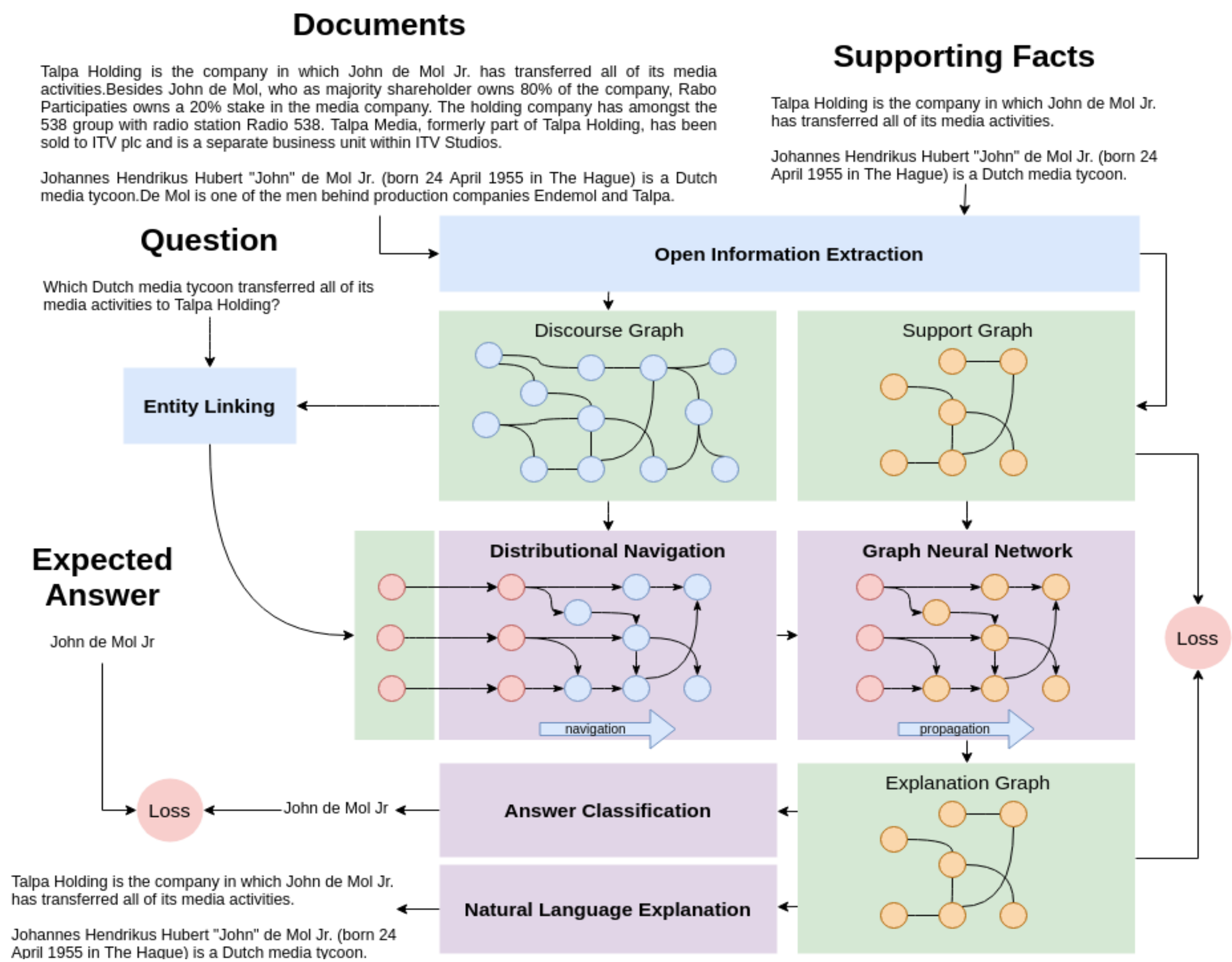
Paragraph B: Toxics Release Inventory

The Toxics Release Inventory (TRI) is a publicly available database containing information on toxic chemical releases and other waste management activities in the United States.

Yang, Zhilin, et al. "HotpotQA: A Dataset for Diverse, Explainable Multi-hop Question Answering." *arXiv preprint arXiv:1809.09600*(2018).

Explainable QA

Neuro-Symbolic = Knowledge Graphs + Gated Graph Neural Networks (GGNN)



Explainable Science QA

Q: A student climbs up a rocky mountain trail in Maine. She sees many small pieces of rock on the path. <u>Which action most likely made the small pieces of rock?</u>

[0]: sand blowing into cracks

[1]: leaves pressing down tightly

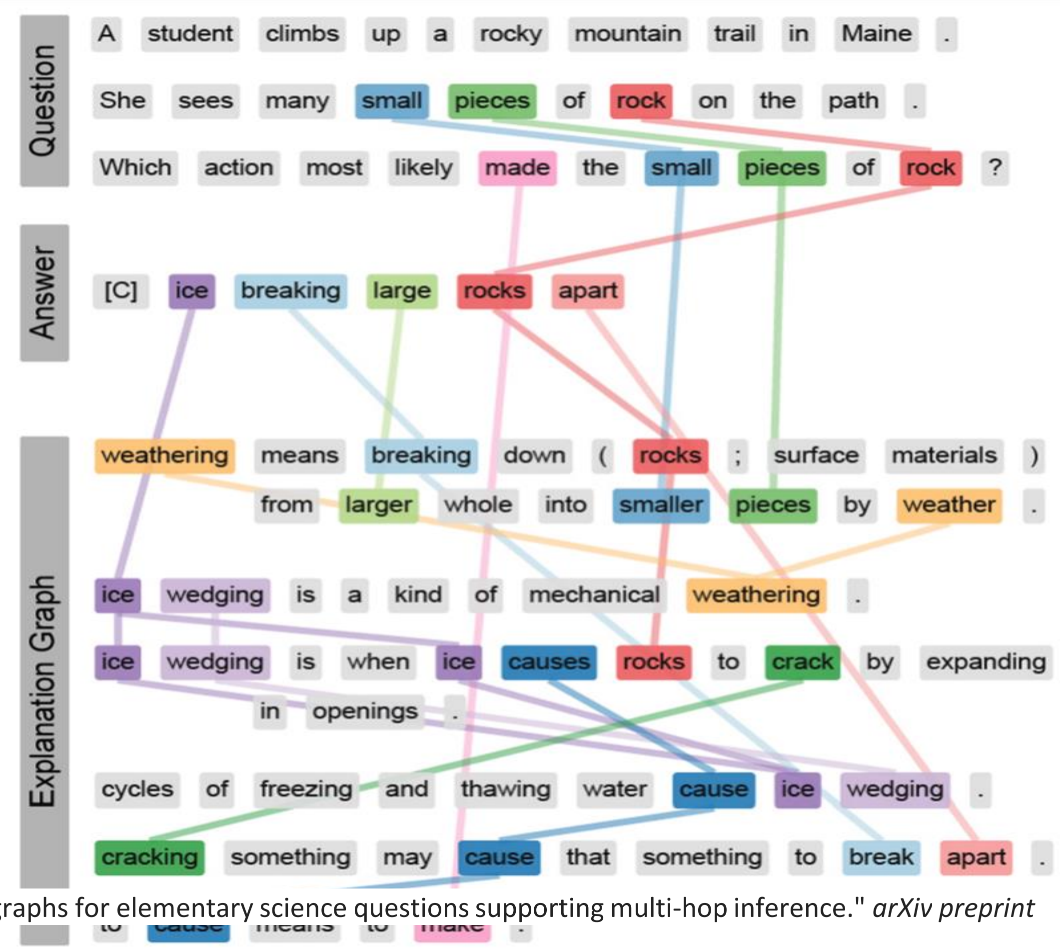
[2]: ice breaking large rocks apart

[3]: shells and bones sticking together

A: ice breaking large rocks apart

Explanation:

- weathering means breaking down (rocks
  weather
- ice wedging is a kind of mechanical weat
- ice wedging is when ice causes rocks to c
- cycles of freezing and thawing water cau
- cracking something may cause that som
- to cause means to make



Jansen, Peter A., et al. "Worldtree: A corpus of explanation graphs for elementary science questions supporting multi-hop inference." *arXiv preprint arXiv:1802.03052* (2018).

LINKQ

Powered by Thomson Reuters

News Feed | Connection | Portfolio

Settings | All | Pinned

3 number of pinned articles

BRIEF-Avio receives EUR 40 mln financing from European Investment Bank <SPA2.MI>

Companies: Avio SPA 80%, Avio SPA 80%, Avio SPA 80%
Topics: Business Finance, Contracts / Business Deals Events: ContactDetails
Industry: Spacecraft Manufacturing
Publication date: Oct 6, 2017 10:24:02 AM

Airbus SE | BAE Systems PLC | Boeing Co

BRIEF-Airtelis orders three H215 Airbus Helicopters<AIR.PA><AIRG.DE>

Companies: Airbus Helicopters SAS 50%, Airbus Helicopters SAS 50%, Airbus Helicopters SAS 50%
Topics: N/A Events: BusinessRelation
Industry: Aerospace & Defense - NEC, Aircraft Parts Manufacturing - NEC
Publication date: Oct 4, 2017 10:46:00 AM

Raytheon Co | BAE Systems PLC | Boeing Co

BRIEF-British Airline Pilots' Association - pilots union calls for investigation into collapse of Monarch Airlines<MONA.UL>

Companies: Monarch Airlines Ltd 80%, Monarch Airlines Ltd 80%, Monarch Airlines Ltd 80%
Topics: Other Events: N/A
Industry: Airlines - NEC
Publication date: Oct 9, 2017 1:55:13 PM

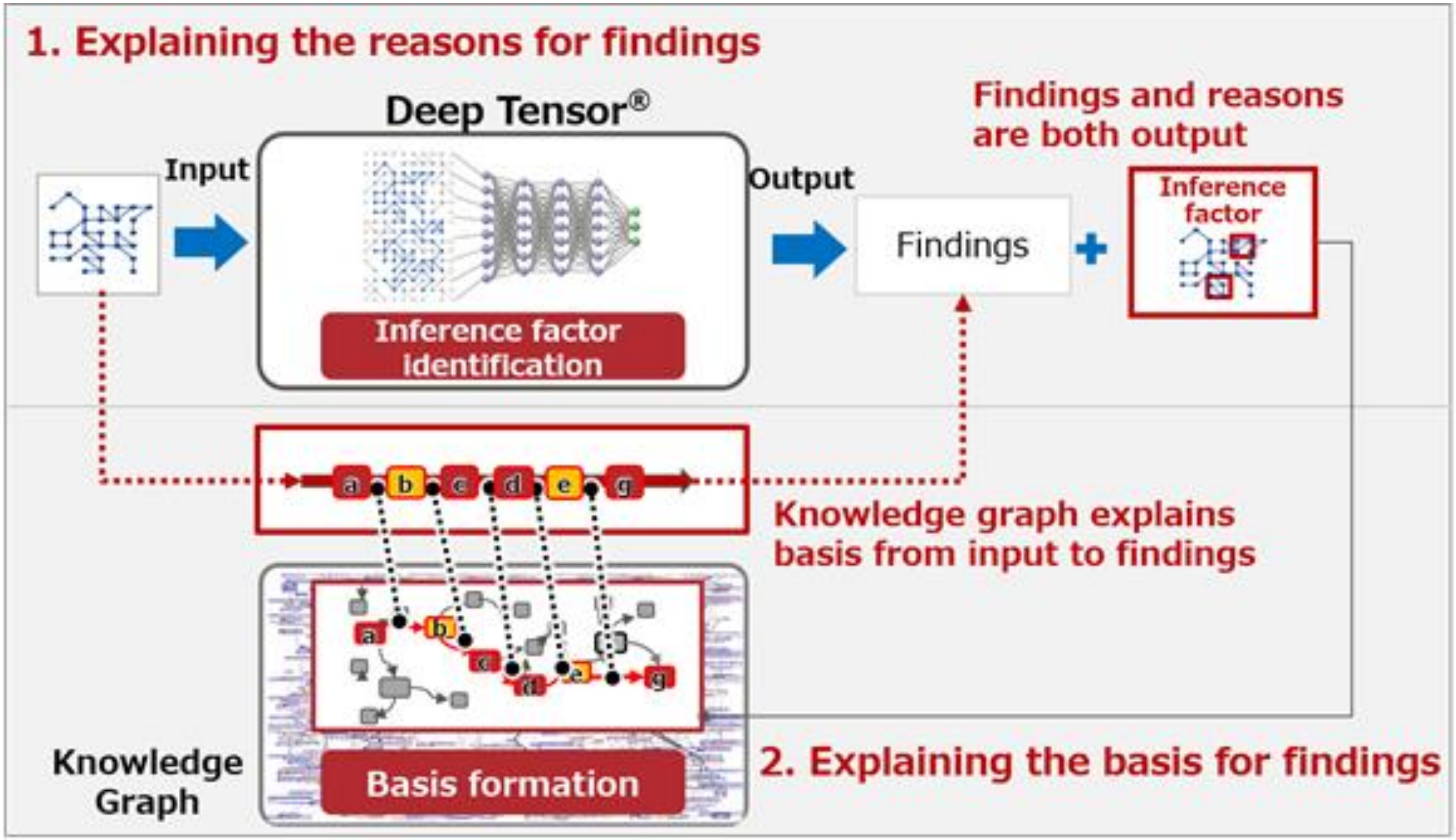
Raytheon Co | BAE Systems PLC | Boeing Co

THOMSON REUTERS

BRIEF-Avio receives EUR 40 mln financing from European Investment Bank <SPA2.MI>

Oct 6 (Reuters) - AVIO SPA <SPA2.MI> * SAYS SIGNED WITH EUROPEAN INVESTMENT BANK CONTRACT FOR EUR 40 MILLION FINANCING Source text for Eikon: [ID:nBIA5D9Ntk] Further company coverage: [SPA2.MI] (Gdynia Newsroom) ((gdynia.newsroom@thomsonreuters.com; +48 58 772 0920 ;))

Explainable Findings
From Tensor Inferences Back to KGs



Summary of Today

- End-to-end overview of this part of the course and its underlying motivation.
- Representation, semantics, learning/inference.
- Representation and NL.
- Dialogue between Statistical and Symbolic approaches.

Next Class

- We will jump directly into Knowledge Representation (beyond FOL).

- Frames, Prototypes, Ontologies.
- Representing more complex NL discourse.

Recommended Reading

Deep Learning:
A Critical Appraisal

Gary Marcus[1]
New York University

https://arxiv.org/ftp/arxiv/papers/1801/1801.00631.pdf