Symbolic AI

Andre Freitas
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.
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 **neuro-symbolic models**.

• Larger palette to build AI Systems.
Michelle Obama (m. 1992)
Barack Obama, Spouse

Michelle Obama - Wikipedia, the free encyclopedia
Michelle LaVaughn Robinson Obama (born January 17, 1964) is an American lawyer, writer, is the wife of the 44th and current President of the United States, Barack Obama. She was born in Chicago, Illinois. She is the first African American woman to be the First Lady of the United States.

U.S. Cities
What is
Toronto??????

What can I help you with?
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 ← t + 1
return action
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
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
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

Problem
Task

KB

Algorithms

Sub-symbolic to Symbolic

Classify, Structure

Perception

Learn
Infer
Program

Answer
Explain
Act
Syllabus

1. Knowledge Representation
2. KB Construction
3. Montague Semantics
4. Semantic Parsing
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
Ceci n’est pas une pipe.
“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
\[ P(r, \theta, \Phi) \]

Cylindrical Coordinates
\[ P(r, \theta, z) \]

\[
\begin{align*}
    ds^2 &= \left(1 - \frac{2m}{r}\right) dt^2 - \frac{1}{\left(1 - \frac{2m}{r}\right)} dr^2 \\
    &\quad - (r)^2 (d\theta^2 + \sin^2(\theta)d\phi^2)
\end{align*}
\]

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.
“The distinctive feature of brains such as the one we own is their uncanny ability to create maps... 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

\[ \text{Formal meaning representation model (lots of data)} \]
\[ + \text{ inference model} \]

This behaves a lot like intelligence!
Semantics

= Formal meaning representation model (lots of data) + inference model

> 2000 years of tradition!

Logics, linguistics, philosophy, cognitive sciences, computer science

This behaves a lot like intelligence!
Building Knowledge Bases
Data

Unstructured Data

Easy to generate

Structure/Semantics

KB Construction

Structured Data

Easy to analyze (computationally)

Consistent

Comparable

Processable
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.
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
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.
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.
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
./.
Barack Obama went with his daughter Malia to the baseball game.
Regularities in Natural Language

Barack Obama went with Malia to the baseball game.

his daughter
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)
- ...

2!
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

\[
S/N \\
\lambda f. f
\]

\[
N \\
\lambda x. flight(x)
\]

\[
PP/NP \\
\lambda y. \lambda x. to(x, y)
\]

\[
NP \\
BOSTON
\]

\[
PP \\
\lambda x. to(x, BOSTON)
\]

\[
N \setminus N \\
\lambda f. \lambda x. f(x) \land to(x, BOSTON)
\]

\[
N \\
\lambda x. flight(x) \land to(x, BOSTON)
\]

\[
S \\
\lambda x. flight(x) \land to(x, BOSTON)
\]
Symbol

Tree

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

Some Goal

Similarity, discrimination

Tree is the name of a set

Tree is a noun

Structure of the Reality

Representation
Symbol

Tree

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

Some Goal

Similarity, discrimination

Extension of the set

Structure of the Reality

Representation
Symbol

Tree

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

Some Goal

Similarity, discrimination

Operating on the Representation

The tallest tree
Symbol

Tree

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

Operating on the Representation

The tallest tree

Definite article: “get me one”

Some Goal

Similarity, discrimination

Representation

Structure of the Reality
Structural similarity, discrimination

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

Symbol: Tree

The tallest tree

Some goal

Similarity, discrimination

Operating on the Representation

Superlative adjective: “top most”

Structure of the Reality

Representation
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
    \[ \text{Grandparent}(x, y) \iff \exists z \text{ Mother}(x, z) \land \text{Mother}(z, y) \lor \exists z \text{ Mother}(x, z) \land \text{Father}(z, y) \lor \exists z \text{ Father}(x, z) \land \text{Mother}(z, y) \lor \exists z \text{ Father}(x, z) \land \text{Father}(z, y) \]
Inductive Logic Programming

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

  \[ \text{Parent}(x, y) \iff [\text{Mother}(x, y) \lor \text{Father}(x, y)] \]
  \[ \text{Grandparent}(x, y) \iff [\exists z \text{ Parent}(x, z) \land \text{Parent}(z, y)] \]

• 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.
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

<table>
<thead>
<tr>
<th>diagram</th>
<th>symbol</th>
<th>name</th>
<th>example</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="equivalence" /></td>
<td>$x \equiv y$</td>
<td>equivalence</td>
<td>$couch \equiv sofa$</td>
</tr>
<tr>
<td><img src="image2" alt="forward entailment" /></td>
<td>$x \sqsubset y$</td>
<td>forward entailment (strict)</td>
<td>$crow \sqsubset bird$</td>
</tr>
<tr>
<td><img src="image3" alt="reverse entailment" /></td>
<td>$x \sqsupset y$</td>
<td>reverse entailment (strict)</td>
<td>$European \sqsupset French$</td>
</tr>
<tr>
<td><img src="image4" alt="negation" /></td>
<td>$x \uparrow y$</td>
<td>negation (exhaustive exclusion)</td>
<td>$human \uparrow nonhuman$</td>
</tr>
<tr>
<td><img src="image5" alt="alternation" /></td>
<td>$x \upharpoonright y$</td>
<td>alternation (non-exhaustive exclusion)</td>
<td>$cat \upharpoonright dog$</td>
</tr>
<tr>
<td><img src="image6" alt="cover" /></td>
<td>$x \square y$</td>
<td>cover (exhaustive non-exclusion)</td>
<td>$animal \square nonhuman$</td>
</tr>
<tr>
<td><img src="image7" alt="independence" /></td>
<td>$x # y$</td>
<td>independence</td>
<td>$hungry # hippo$</td>
</tr>
</tbody>
</table>
Neuro-Symbolic Models
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

Statistical vs Symbolic AI Systems

<table>
<thead>
<tr>
<th></th>
<th>Statistical</th>
<th>Symbolic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explainability</td>
<td>Hard</td>
<td>Easy</td>
</tr>
<tr>
<td>Generalizing algebraic operations</td>
<td>Hard</td>
<td>Easy</td>
</tr>
<tr>
<td>Robustness to noise</td>
<td>Easy</td>
<td>Hard</td>
</tr>
<tr>
<td>Robustness to ambiguity</td>
<td>Easy</td>
<td>Hard</td>
</tr>
<tr>
<td>Robustness to mislabeling</td>
<td>Easy</td>
<td>Hard</td>
</tr>
</tbody>
</table>
It is possible for systems to combine statistical perceptual with conceptual interpretable reasoning!

<table>
<thead>
<tr>
<th></th>
<th>Deep Learning</th>
<th>Symbolic Program Synthesis</th>
<th>δILP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Robust to noise</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td>Can learn from non-symbolic data</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td>Data efficient</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Interpretable</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

Gated Graph Neural Networks

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

Nodes $v \in V$ take unique values from $1, \ldots, |V|$, and edges are pairs $e = (v, v_0) \in V \times V$

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 $o_v = g(h_v, l_v)$ maps from node representations and corresponding labels to an output $ov$ for each $v \in V$

The propagation model is similar to an LSTM. Each node in the graph $v$ has a hidden state representation $h(t)\nu$ that is updated at every time step $t$. The computation starts at $t = 0$ with initial hidden states $x_v$ 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 gated update module.

Building Neuro-Symbolic Systems
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.
Explainable QA

Neuro-Symbolic = Knowledge Graphs + Gated Graph Neural Networks (GGNN)
Q: A student climbs up a rocky mountain trail in Maine. She sees many small pieces of rock on the path. **Which action most likely made the small pieces of rock?**

[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; surface materials) from larger whole into smaller pieces by weather
- ice wedging is a kind of mechanical weathering
- ice wedging is when ice causes rocks to crack by expanding
- cycles of freezing and thawing water cause ice wedging
- cracking something may cause that something to break apart
- to cause means to make

BRIEF-Avio receives EUR 40 mn 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

BRIEF-AirTelis orders three H215 Airbus Helicopters <AIR.PA> <AIRG.DE>
Companies: Airbus Helicopters SAS 60%, Airbus Helicopters SAS 60%, Airbus Helicopters SAS 60%
Topics: N/A Events: BusinessRelation
Industry: Aerospace & Defense - NEC, Aircraft Parts Manufacturing - NEC
Publication date: Oct 6, 2017 10:46:00 AM

BRIEF-British Airline Pilots' Association - pilots union calls for investigation into collapse of Monarch Airlines <MONA.U>.
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
BRIEF: Avio receives EUR 40 mn 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:nBIA5D9N1k] 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
New York University