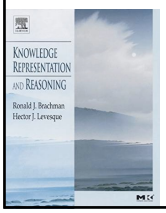
THE UNIVERSITY of EDINBURGH

INFR11215 Knowledge Graphs

Introduction

Jeff Z. Pan

http://knowledge-representation.org/j.z.pan/



[Reading: Brachman and Levesque, chapters 1 and 8]

1

THE UNIVERSITY of EDINBURGH

Lecture Outline

- Knowledge Representation and Reasoning
  - What and Why
  - A Bit of History
- This course
  - Course Structure
  - Reading List

2

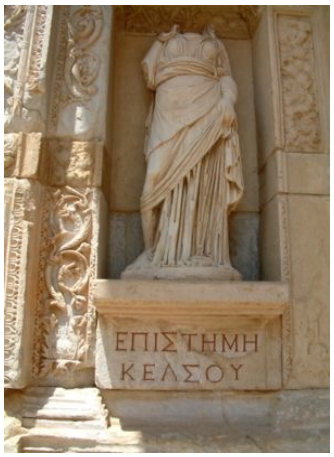
THE UNIVERSITY of EDINBURGH

What is Your Definition of Knowledge?

3

THE UNIVERSITY of EDINBURGH

Knowledge



- What is knowledge?
  - Something that is known
  - Verified belief
  - Likely to be Structured information
  - About certain aspects of the (real) world

4

THE UNIVERSITY of EDINBURGH

Example

- For example, take this list:
  - (31 28 24 23 16)
- This mere data becomes knowledge
  - if you knew that these were temperatures of five successive days
  - Temp(day1, 31), …, Temp(day5, 16)
- Knowing this allows you to answer questions about the data
  - Such as “Is it getting cooler?”

5

THE UNIVERSITY of EDINBURGH

Knowledge Representation

[Temp(day1, 31), …, Temp(day5, 16)]

- KR is a surrogate
  - A substitute for the thing itself
- Level 1: Symbol
  - 8 vs VIII
  - New York vs Big Apple
- Level 2: Dependency among symbols
  - 1+2=3
  - Eucalypt is a kind of Plant
- Knowledge Rrepresentation is central to AI

6

THE UNIVERSITY of EDINBURGH

Representation Matters

CXLII + DCCCLVIII =?

142+ 858 =?

- Dialog between two heavy smokers and a priest
  - Heavy Smoker 1: Can I smoke when I pray?
  - Priest: No.
  - Heavy Smoker 2: Can I pray when I smoke?
  - Priest: Yes.

7

THE UNIVERSITY of EDINBURGH

Explicit vs Implicit Knowledge

- Explicit knowledge
  - Unstructured knowledge: texts, images, videos
  - Structured knowledge: knowledge bases, databases
- Implicit knowledge
  - Parametric knowledge: such as those in Large Language Models (LLMs)
  - Other implicit knowledge: such as skills depending on human sensers and memory units

8

THE UNIVERSITY of EDINBURGH

Why Knowledge Representation

- *Knowledge is not the destination*
  - We want actions
  - We want skills to perform the right actions in the right place at the right time

Photo Credit: serendipitymommy.com

- Wouldn't it be better to "complie out" the KB and distribute the knowledge to the precedures that needs it?

9

THE UNIVERSITY of EDINBURGH

Nature of Expertise: Chess

EXPERIMENT 1 (Simon & Chase, 1973)

Experts & Novices were both shown a chess board of an actual game. Both groups were shown the board for the same period. Both groups were asked to recall what pieces they could remember.

Result: The Experts were able to recall significantly more pieces than the Novices.

EXPERIMENT 2

Both experts & novices were shown a Chess board where pieces had been placed randomly, for the same period.

Result: Both groups recalled similar numbers of pieces.

CONCLUSION: EXPERTS save their KNOWLEDGE as CHUNKS (i.e. meaningful patterns).

[Credit: D Sleenman]

10

THE UNIVERSITY of EDINBURGH

Nature of Expertise: Programming

- Expert & novice programmers shown:
  - Actual code {Experts recall much more than novices: that is they see structures like while loops whereas novices see reserved words like while}
  - Programs where the lines have been randomly presented {Experts & novices recall comparable amounts}

[Credit: D Sleenman]

THE UNIVERSITY of EDINBURGH

NATURE of EXPERTISE: SUMMARY

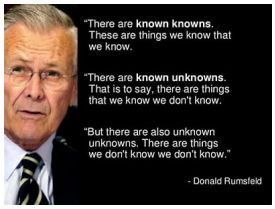
- Expert's knowledge is,
  - in a sense, "compiled" to enhance performance, and
  - this *may* impede explanation.
  - real experts can "extend" their knowledge to truly novel tasks.
- Expertise (both in intellectual and physical skills)
  - takes many years to acquire and compile and
  - needs regular usage to maintain
- (Even) experts do not always/ often systematically investigate large solution spaces.
  - So it is possible, in some circumstances, for a Knowledge based System to outperform an "expert". (Based on a quote from Bruce Buchanan.)

https://www.cell.com/current-biology/pdf/S0960-9822(14)00413-8.pdf

[Credit: D Sleenman]

THE UNIVERSITY of EDINBURGH

Known & Unknowns



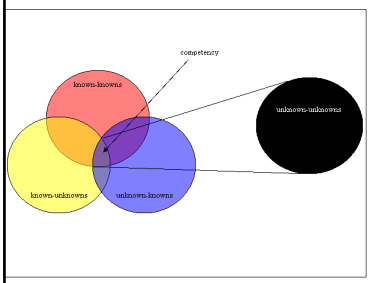
[Picture Credit: www.goodreads.com]

- There are known knowns
  - these are things we know we know.
- We also know there are known unknowns
  - this is to say we know there are some things we do not know.
- There are also unknown knowns
  - these are things we do not know we know
- But there are also unknown unknowns
  - the ones we don't know we don't know.

13

THE UNIVERSITY of EDINBURGH

Known & Unknowns (II)



[Picture Credit: www.akkamsrazor.com]

- Experts have vast amounts of knowledge to share
  - yet they understand there are open questions in the domain
- They don't know all that they know and use
  - Experts have a lot of tacit knowledge
  - Tacit knowledge is hard (impossible) to describe.
- Each expert doesn't know everything

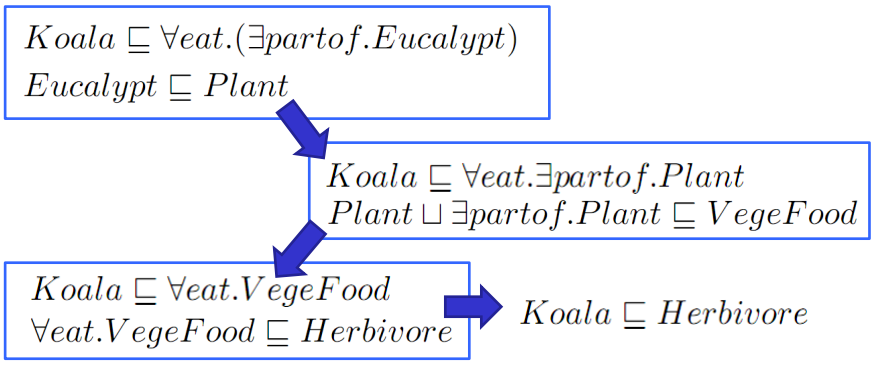
14

THE UNIVERSITY of EDINBURGH

Reasoning

- Turning something from we don't know that we know into we know that we know

(*please don't worry about the syntax of the statements below for now*)



15

THE UNIVERSITY of EDINBURGH

A bit of history: Semantic Nets

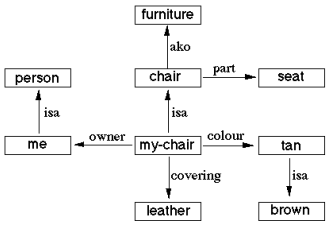
- Proposed by Quillian (1968) to analysis the meaning of words in sentences
  - Later applied to KR
- Basic notations:
  - Nodes: to represent objects, concepts, or situations
  - Arcs: to represent relationships

16

THE UNIVERSITY of EDINBURGH

Pros/Cons of Semantic Nets

- Easy to follow hierarchy
- Easy to trace association
- Flexible



- No well defined syntax
- No formal semantics
- Not expressive enough to define meaning of labels
- Inefficient

THE UNIVERSITY of EDINBURGH

Frames

- A frame [Minsky 1981] represents knowledge
  - About a narrow subject
  - That has default inheritance
- Example:

```
(Elephant
   <:IS-A Mammal>
   < :Colour grey>
   <:EarSize  large> ...)
(raja:
 <:INSTANCE-OF Elephant>
 <:EarSize small> ...)
(RoyalElephant
  <:IS-A Elephant>
  <Colour white> ...)
```

THE UNIVERSITY of EDINBURGH

Procedures in Frames

- Frames allow procedures called demons to be attached to their slots (properties)
  - **IF-ADDED:** triggered when a new value is put into a slot
  - **IF-REMOVED**: triggered when a value is removed from a slot
  - **IF-REPLACED:** triggered when a slot value is replaced
  - **IF-NEEDED:** triggered when there is no value present in an instance frame and a value must be computed from a generic frame

19

THE UNIVERSITY of EDINBURGH

Example: Procedures in Frames

```
(Person
  <:age
    [range      1..150]
    [IF-NEEDED
            [ask("What  is the age of ", name of this person);]]
    [IF-ADDED
            [if new value > 150 then
                    print("Your ", this slot, " is too high!");]]
)
```


20

THE UNIVERSITY of EDINBURGH

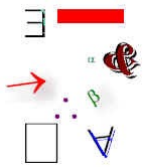
Pros/Cons of Frames

- Easy to include default information and detect missing values
- Easy to create specialised procedures

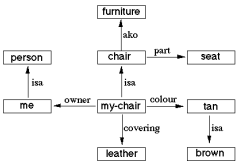
- Difficult to program
- Difficult to provide a logical semantics
- Difficult for inference

THE UNIVERSITY of EDINBURGH

First Order Logic



- Allow to use quantifiers in sentences
  - "for all" (∀), "exists" (∃)
  - Makes sentences more precise



- Example:
  - Chair is a sub-class of furniture
    - ∀x (chair(x) → furniture(x))
  - A chair has a part seat
    - ∀x (chair(x) → ∃ y (part(x,y) ∧ seat(y))

THE UNIVERSITY of EDINBURGH

Pros/Cons of FOL

- Expressive
- Clear formal semantics

- Syntax too complex
- No support for uncertainty
- Not deci

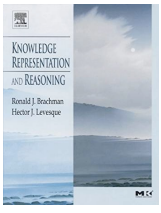
Notes:
- Being **decidable** means there exists an algorithm returning all and only the correct answers in finite time.

23

THE UNIVERSITY of EDINBURGH

KL-ONE

- Ron Brachman: Use FOL-based formal semantics, but limit the expressive power
- Formalising Semantic Nets
  - Three kinds of vocabulary
    - Concepts, properties and objects
  - Non-graphic Syntax
  - Formal semantics for built-in relationships
    - Sub-Class-Of
    - Instance-of
  - Provide constructors to define concepts



24

THE UNIVERSITY of EDINBURGH

Example: KL-ONE

- Example:
  - Chair is a sub-class of furniture
    - Chair $\sqsubseteq$ Furniture
  - A chair has a part seat
    - Chair $\sqsubseteq \exists$ *part*.Seat

25

THE UNIVERSITY of EDINBURGH

Pros/Cons of KL-ONE like languages

- Well defined and simplified syntax
- Clear formal semantics
- Good balance between expressive power and decidability

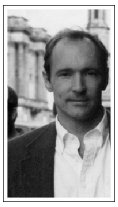
[Later known as **Description Logics**, and used as the Semantic Web standard **ontology language**]

- Not as expressive as FOL
- No support for procedures

26

THE UNIVERSITY of EDINBURGH

How to Scale it Up

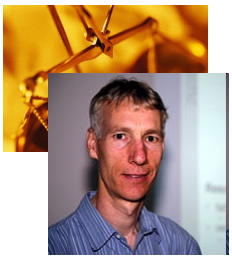


- Sir Tim Berners-Lee: Father of the Web
  - Allow people to share knowledge
  - To work together better
- Evolving Web
  - Version 1 (Web): unstructured knowledge
  - Version 2 (Semantic Web): structured knowledge
- Allow programs to handle the shared knowledge and data
  - In a meaningful way
  - To help people analysis and manage activities
  - Standards: RDF, OWL, SPARQL, ...

27

THE UNIVERSITY of EDINBURGH

Really, how to scale it up



- However, reasoning is not trivial
  - Decidability
  - Tractability
- Breakthrough around 2000-2007
  - Highly optimisation algorithms for ontology reasoners
  - By Prof. Ian Horrocks (Oxford U.) and colleagues
    - FaCT reasoner (Ian Horrocks)
    - FaCT-DG reasoner (Jeff Pan and Ian Horrocks)
    - FaCT++ reasoner (Dmitry Tsarkov and Ian Horrocks)
- Some later breakthrough (2008-2016)
  - TrOWL: Approximate reasoning (Jeff Pan)
  - ...

28

THE UNIVERSITY of EDINBURGH

Knowledge Graph

- A knowledge graph is a set of interconnected typed entities and their attributes
  - with a schema defining related vocabulary
  - Entities are semantically rich objects for disambiguation and interoperability
- Used by Google from 2012
  - allowing users to search for things, people or places (1 billion entities, 70 billion facts)
- International standards: RDF, OWL, SPARQL
- KG = Data sub-graph + schema sub-graph

Freebase

DBpedia

yago select knowledge

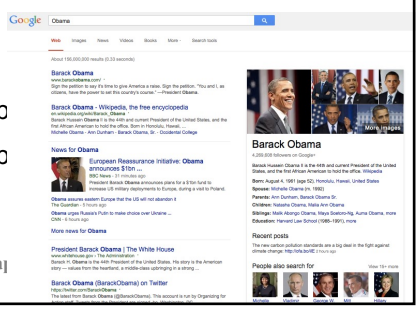
29

THE UNIVERSITY of EDINBURGH

Data Sub-Graph

DBpedia

- DBpedia data sub-graph is represented in Resource Description Framework (RDF, W3C standard for Knowledge Graph)
  - [dbr:Barack_Obama rdf:type dbo:President .]
  - [dbr:Barack_Obama dbo:birthPlace dbr:Hawaii .]
  - [dbr:Barack_Obama dbo:spouse dbr:Michelle_Obama .]
- DL syntax
  - dbr:Barack_Obama : dbo:President
  - (dbr:Barack_Obama, dbr:Hawaii) : db
  - (dbr:Barack_Obama, dbr:Michelle_Ob

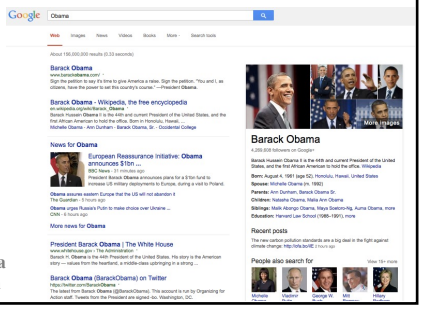


30

THE UNIVERSITY of EDINBURGH

Schema Sub-Graph

- DBpedia data schema-graph is represented in an OWL (Web Ontology Language, W3C standard for Knowledge Graph schema, based on Description Logics) ontology http://dbpedia.org/ontology/
  - dbo:President ⊑ dbo:Politician
  - dbo:spouse ⊑ dbo:spouse⁻



THE UNIVERSITY of EDINBURGH

Reasoning with Knowledge Graph

- Example reasoning rule:
- C ⊑ D, [b rdf:type C .] => [b rdf:type D .]
  - from dbo:President ⊑ dbo:Politician and [dbr:Barack_Obama rdf:type dbo:President .]
  - we have [dbr:Barack_Obama rdf:type dbo:Politician .]
- If we have some extra schema knowledge, dbo:President ⊑ ¬dbo:Mayor
  - then we know that the following fact is not true
  - [dbr:Barack_Obama rdf:type dbo:Mayor .]

THE UNIVERSITY of EDINBURGH

Hybrid Representation

[Picture Credit: pngitem.com]

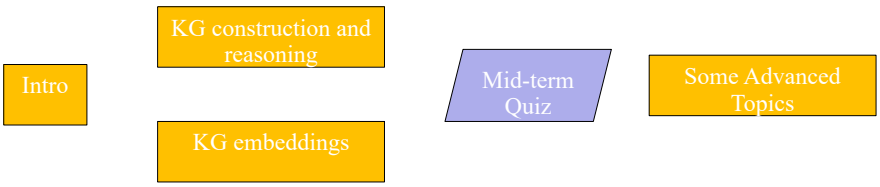
- Highlights of KGs:
  - Combination of logical reasoning and graph based learning
  - Friendly to human, explainable logical reasoning, easy to reuse and manage
- Highlights of LLMs:
  - Good coverage of commonsense knowledge
  - Support many downstream tasks
- Cannot replace one another

https://arxiv.org/pdf/2308.06374.pdf

33

THE UNIVERSITY of EDINBURGH

Course Roadmap



34

THE UNIVERSITY of EDINBURGH

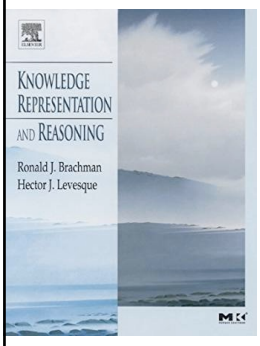
This Course

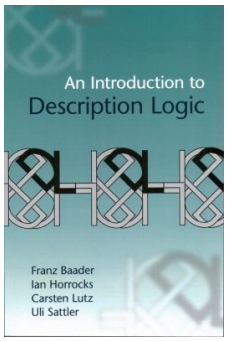
- Lectures
  - Tuesday
  - Friday
- Tutorials
  - Weeks 2, 4, 8, 10
- Lab Practicals
  - Weeks 4,5, 8, 9, 10

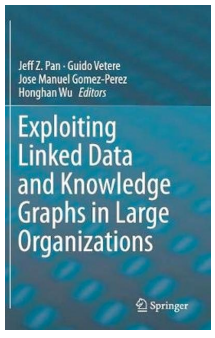
35

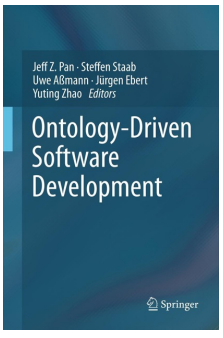
THE UNIVERSITY of EDINBURGH

Reading List









https://opencourse.inf.ed.ac.uk/kg/resource-list

36

THE UNIVERSITY of EDINBURGH

Lecture Outline

- Motivation: Knowledge is central to intelligence
- Introduction: What is Knowledge Representation and Reasoning (KRR)
- The modern standard approach to KRR: Knowledge Graph
- Exercises (Next time we will explain how to express complex facts)
  - Some rich men love Jenny
  - All rich men love Jenny
  - Jenny loves both John and Jim
  - Jenny loves either John or Jim

37