Knowledge Representation and Reasoning

Artificial Intelligence Challenge / Introduction to Artificial Intelligence

ICM 2A + M1 CPS²

8th March 2023

antoine.zimmermann@emse.fr

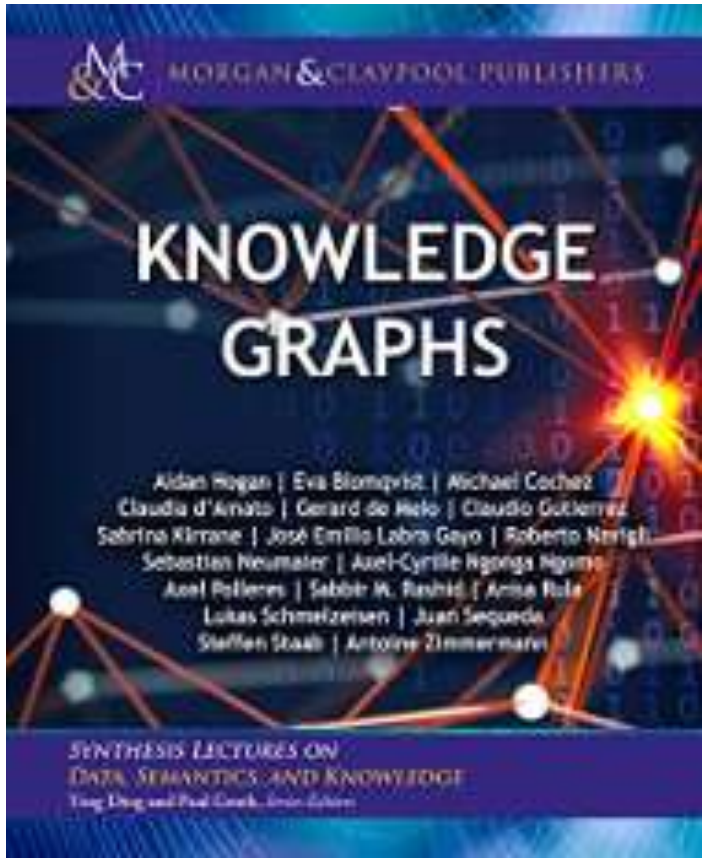
Organisation of the course

- One Web page where I put all the resources for this course:

  https://www.emse.fr/~zimmermann/Teaching/KRR/

- 6 sessions, including an exam, mixing *lectures*, *tutorials*, and *lab sessions*
- Focus on knowledge **representation** more than *reasoning techniques*
- Evaluation based on homework in groups and online MCQs
- I hope classwork can be interactive: student <-> student interactions as well as student <-> teacher interaction
- Bring your computers. You are **encouraged** to look up things on the Web when I'm presenting, but **no games**, **no social networks**, **no chat** during classes 😈 and **keep your attention on what I'm saying**

Book I'll use for this course



- Hogan et al. (2021): *Knowledge Graphs*. Morgan & Claypool.
- https://kgbook.org/

What is *reasoning*?

- There are different forms of reasoning
  - Inductive reasoning
    - Build general theories (knowledge) from particular observations (facts)

    e.g. *If it has wheels, doors, seats, windows, engine, it must be a car.*
  - Deductive reasoning
    - Use general theories (knowledge) on particular cases to infer new facts

    e.g. *All cars are vehicles. Emmanuel Macron's DS 7 Crossback is a car. Therefore, Macron's DS 7 Crossback is a vehicle.*
  - Abductive reasoning
    - Use general theories and incomplete observations to get most plausible conclusion

    e.g. *My car is not in my garage. Most likely, my wife must still be at her work place.*

    e.g. *My car is not in my garage and my wife is at home. It may have been stolen!*

Interactions between the types of reasoning

- Inductive reasoning can be used to devise the knowledge used later in deductive reasoning
- Deductive reasoning can be used to avoid inducing inconsistent knowledge or to validate inductive hypothesis (e.g., the Chinese hypothesis)
- Deductive reasoning can be used to check if an abductive reasoning is plausible
- Abduction can be used in combination with inductive reasoning to only induce the more plausible facts
- Etc.

Inductive reasoning vs. deductive reasoning (1)

- Playing Go requires knowledge of the rules of the game
- From knowledge of the rules, one can decide if a stone placement is allowed or dissallowed (*deductive reasoning*)
- One *could* learn Go from mere observation of games (*inductive reasoning*), but they would have a chance of making illegal placement, or may not attempt placements that are in fact legal
- It is **relatively easy** to formalise the rules of Go (with few corner cases that are trickier)
- It is **very difficult** to ensure that we know all legal/illegal placements by observation. It is even more difficult to define scoring this way

Inductive reasoning vs. deductive reasoning (2)

- Winning a game of Go requires a lot of experience
- We do not know how to define general knowledge of the game such that the next move logically follows from our knowledge
- Playing good Go is better achieved by inductive reasoning from observation of a lot of games
- **But** players still need to know exactly the rules of the game to avoid making an illegal move

Inductive reasoning and the Chinese hypothesis (1)

For a natural number $n > 1$, is the number $2^n - 2$ divisible by n ?

$2^2 - 2 = 2 = 1 \times 2$

$2^3 - 2 = 6 = 2 \times 3$

$2^4 - 2 = 14 = 2 \times 7$

$2^5 - 2 = 30 = 6 \times 5$

$2^6 - 2 = 62 = 2 \times 31$

$2^7 - 2 = 126 = 18 \times 7$

$2^8 - 2 = 254 = 2 \times 127$

$2^9 - 2 = 510 = 2 \times 3 \times 5 \times 17$

$2^{10} - 2 = 1\,022 = 2 \times 7 \times 73$

$2^{11} - 2 = 2\,046 = 186 \times 11$

$2^{12} - 2 = 4\,094 = 2 \times 23 \times 89$

$2^{13} - 2 = 8\,190 = 630 \times 13$

$2^{14} - 2 = 16\,382 = 2 \times 8\,191$

$2^{15} - 2 = 32\,766 = 2 \times 3 \times 43 \times 127$

$2^{16} - 2 = 65\,534 = 2 \times 7 \times 31 \times 151$

$2^{17} - 2 = 131\,070 = 7\,710 \times 17$

$2^{18} - 2 = 262\,142 = 2 \times 131\,071$

$2^{19} - 2 = 524\,286 = 27\,594 \times 19$

$2^{20} - 2 = 1\,048\,574 = 2 \times 524\,287$

Inductive reasoning and the Chinese hypothesis (2)

For a natural number $n > 1$, is the number $2^n - 2$ divisible by n ?

21 ✗
22 ✗
23 ✓
24 ✗
26 ✗
27 ✗
28 ✗
29 ✓

30 ✗
31 ✓
32 ✗
33 ✗
34 ✗
35 ✗
36 ✗
37 ✓
38 ✗
39 ✗
40 ✗

41 ✓
42 ✗
43 ✓
44 ✗
45 ✗
46 ✗
43 ✓
48 ✗
49 ✗

Inductive reasoning and the Chinese hypothesis (3)

- Is it the case that:

  for all prime number *n*, the number $2^n - 2$ is divisible by *n*

???

Inductive reasoning and the Chinese hypothesis (3)

- Is it the case that:

  for all prime number *n*, the number $2^n - 2$ is divisible by *n*

???

Yes it is!

Beware inductive reasoning

Machine learning is not good for everything

Inductive reasoning and the Chinese hypothesis (4)

- But, wait:

Is the number $2^{341} - 2$ divisible by 341

???

Inductive reasoning and the Chinese hypothesis (4)

- But, wait:

Is the number $2^{341} - 2$ divisible by 341

???

Yes it is!

…but $341 = 11 \times 31$ is not a prime

Inductive reasoning vs. deductive reasoning (3)

- Deductive reasoning is crucial in situation when the **risk of failure** of inductive reasoning overcome the **benefit of probable success**
- Inductive reasoning is particularly good for tasks recognising a situation based on perception

e.g., *based on the intensity of light signals and my experience of seeing urban areas, I can say that the traffic light on the car lane of my car is red*

- Deductive reasoning is particularly good for tasks that require strict application of rules

e.g., *considering that I drive a car on a car lane where the traffic light is red, and I do not have a priority vehicle (ambulance, etc), I must stop*

What is this course about?

- Traditionally, KR&R (and this course) is about the techniques to computerise **deductive reasoning**. In particular:
  - How can knowledge can be represented in machines to allow the computation of deduction? → *symbolic knowledge representation*
  - How to interpret the digital representation in terms of what it *means* → *formal semantics*
  - How to manipulate the digital representation of knowledge to automate reasoning? → *data structures and algorithms*
  - How to use these techniques adequately in concrete use cases? → *knowledge engineering*

Requirements

- The foundations of KR&R are strongly related to logics
- Connections with EMSE 1st year course on *Introduction to logic*
  - The booklet of the logic course is available online (in French) at https://www.emse.fr/~zimmermann/Teaching/Logique/Livret/
- A good understanding of the basis of propositional and first order logics is *preferable*
- Basic programming information system skills are useful.

Where is KR&R used?

These products embed, among other things, a knowledge base, and can do more or less advanced reasoning:

- virtual assistants (Alexa, Cortana, Siri, etc.)
- Wolfram Alpha
- Main search engines (the "Google Knowledge Graph" + schema.org)
  →consequently, web sites that want to be better indexed by search engines
- Enterprise Knowledge Graphs (Airbnb, Amazon, eBay, IBM, LinkedIn, Microsoft, Uber, …)
- Library data management systems (LoC, BNF, BNL, DNB, …)
- Wikidata

Highlight on *Knowledge Graphs*

- Knowledge Graph is a hot topic in AI that lies at the convergence of:
  - Knowledge Representation and Reasoning
  - Data management and processing (and Big Data)
  - Data analytics
  - Machine Learning (and graph embeddings)
  - The World Wide Web (and the Semantic Web)
  - (Linked) Open Data, FAIR scientific data (**F**indable, **A**ccessible, **I**nteroperable, **R**eusable data)

MOOCS (in case Covid-19 keeps you at home)

- *CS 520: Knowledge Graphs.* (2021 course) https://web.stanford.edu/~vinayc/kg/2021/
- *AI: Knowledge Representation and Reasoning*. https://www.classcentral.com/course/swayam-ai-knowledge-representation-and-reasoning-7958
- *Knowledge Engineering with Semantic Web Technologies*. https://open.hpi.de/courses/semanticweb2015
- *Knowledge Graphs*. (2020 course) https://open.hpi.de/courses/knowledgegraphs2020
- *CS 227: Knowledge Representation and Reasoning*. https://web.stanford.edu/class/cs227/ (no videos, just written resources)

+ There are more links to online resources on my web space

Practical work: eCommerce knowledge graph

- What an intelligent agent needs to **know** to buy things online?
- What is it that I need?
- Where can I find it?
- How do I know something is what I need?
- How do I know something is for sale and I can buy it?
- How can I buy it?
- When I buy it, how to I get it?