Ontologies and Classification: The Unavoidable Interplay Between Human Reasoning and Machine Reasoning

SESSION:
Information Science and Knowledge Organization in Motion:
the new data science approach and its underlying technologies

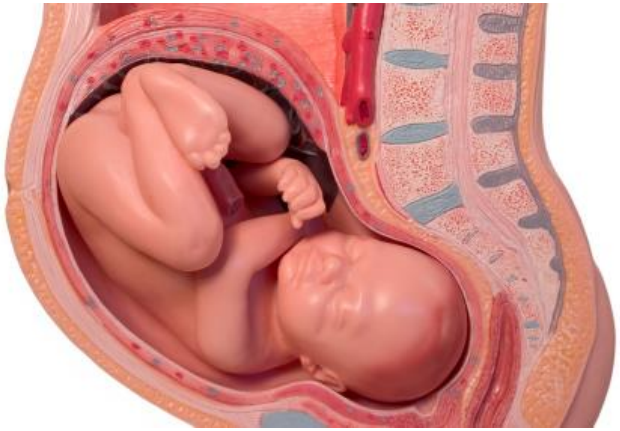
Prof. Mauricio B. Almeida
Federal University of *Minas Gerais*

Florence, Italy, May 2019

Motivation

- Within digital data environments, there are no shelves, no physical constraints that demand the organization for which classification schemes were created
- … no prior scheme can say in advance what a user needs
- … to what extent automatic reasoning can replace human reasoning in classification, an essential activity in LIS

Motivation



"When does a human being begin to exist?"

Smith, B., Brogaard, B. Sixteen Days. (2003). J Med Philos. 2003 Feb;28(1):45-78.

When does an organism has features that qualify it to belong to the category of human beings?

Facebook labels declaration of independence as 'hate speech'

The website told a local newspaper they violated its community guidelines by posting the original document

The New York Times Magazine

FEATURE

The Great A.I. Awakening

How Google used artificial intelligence to transform Google Translate, one of its more popular services — and how machine learning is poised to reinvent computing itself.

DOUGLAS HOFSTADTE

HOME / BROWSE JOURNALS & BOOKS / ...NTATION / VOLUME 68, ISSUE 3 / IS CLASSIFICATION NECESSARY AFTER GOOGLE?

Is classification necessary after Google?

Author(s): Birger Hjørland (Royal School of Library and Information Science, Copenhagen, Denmark)

Approach

Main steps of our approach:

1. We explain human reasoning including classification / categorization
2. We analyse current main approaches to machine reasoning
3. We check possibilities of replacing human reasoning in classification

Our Premise...

"Human beings use categories to think and to speak, and also algorithms of Artificial Intelligence need some kind of category schema to properly run."

source: Saracevic (1996, p.1).

Outline

I. Human reasoning – an overview
   - Theories of Reasoning
   - Classification
II. Machine reasoning
   - Reasoning in ontologies
   - Reasoning in machine learning
III. Discussion
IV. Final Remarks

PART I- Human Reasoning

Topic 1- Theories of reasoning

Topic 2- Classification and Categorization

I-1 Theories of reasoning

1. Ruled-based approach (also, sentential-rule approach)
2. Mental model approach
3. Relational complexity approach

Theories of reasoning

- **Ruled-Based approach**:
  - It is based on formal systems
  - It draws conclusions from sentences using logical connectives and quantifiers
  - E.g. if anyone knows that John thinks all parties are boring, one can see that John will not be present at the Mary´s party tonight
- Issues:
  - Great number of models produced: 64 possible valid syllogisms for a 2 premises and 1 conclusion
  - Having so many models in memory, some are forgotten or neglected

Theories of reasoning

- **Mental model approach**
  - To reason, people create models of reality
  - Then, people check the conclusions taken against the models
- Issues:
  - Models are conceptualized in different ways
  - There are no criteria to choose the better way

Theories of reasoning

- **Probabilistic approach**
  - People make mistakes in reasoning tasks
  - It uses conditional probability rather than logics
- Issues
  - It does not have a reasonable explanation to human reasoning

I-2 Classification and Categorization

- It is impossible to think without categories
- So many questions:
  - What is the relationship between classes and the world itself?
  - Is there an unique top-most category or several ones?
  - How to distinguish one category from another?
- Main historical schools of thought (Philosophy)
  - Essentialism
  - Cluster analysis
  - Historical classification

PART II- Machine Reasoning

Topic 1- Logical Inferences with Ontologies

Topic 2- Probability Inferences with Machine Learning

II-1 Logical Inferences in Ontologies

- Representation languages composed by logical sentences
- They are assertives, stating facts about the world
- They have a syntax and a semantic
- Inference engines: automatic reasoning in ontologies
- Description Logics: creating of definitions for classes
- **Classification** is the process of:
  - checking whether an instance belongs to a class
  - checking whether a class is a subset of other classes

Logical Inferences in Ontologies

The web we all know

Omeprazole

From Wikipedia, the free encyclopedia

Omeprazole (INN) /oʊˈmɛprəzoʊl/ (Prilosec and generics such as losepine) is a proton pump inhibitor used in the treatment of dyspepsia, peptic ulcer disease, gastroesophageal reflux disease, laryngopharyngeal reflux, and Zollinger–Ellison syndrome.

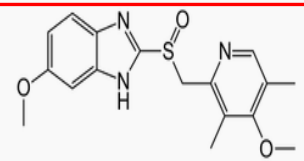
Omeprazole is one of the most widely prescribed drugs internationally and is available over-the-counter in some countries.

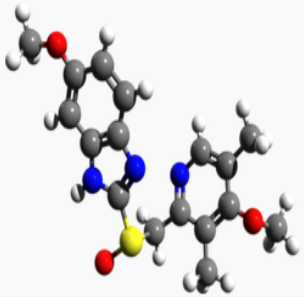
It is on the World Health Organization's List of Essential Medicines, a list of the most important medications needed in a basic health system.[3]

Contents [hide]

1 Medical uses
- 1.1 Gastroesophageal reflux disease
- 1.2 Peptic ulcers
- 1.3 Erosive esophagitis
- 1.4 Zollinger Ellison syndrome

Omeprazole





Systematic (IUPAC) name

6-methoxy-2-((4-methoxy-3,5-dimethylpyridin-2-yl) methylsulfinyl)-1*H*-benzo[*d*]imidazole

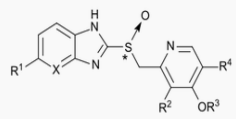
Proton-pump inhibitor

From Wikipedia, the free encyclopedia
(Redirected from Proton pump inhibitor)

Proton-pump inhibitors (**PPIs**) are a group of drugs whose main action is a pronounced and long-lasting reduction of gastric acid production. They are the most potent inhibitors of acid secretion available. The group followed and has largely superseded another group of pharmaceuticals with similar effects, but a different mode of action, called H_2-receptor antagonists. These drugs are among the most widely sold drugs in the world, and are generally considered effective.[1] The vast majority of these drugs are benzimidazole derivatives, but promising new research indicates the imidazopyridine derivatives may be a more effective means of treatment.[2] High dose or long-term use of PPIs carries a possible increased risk of bone fractures.[3]

Proton-pump inhibitor

Drug class

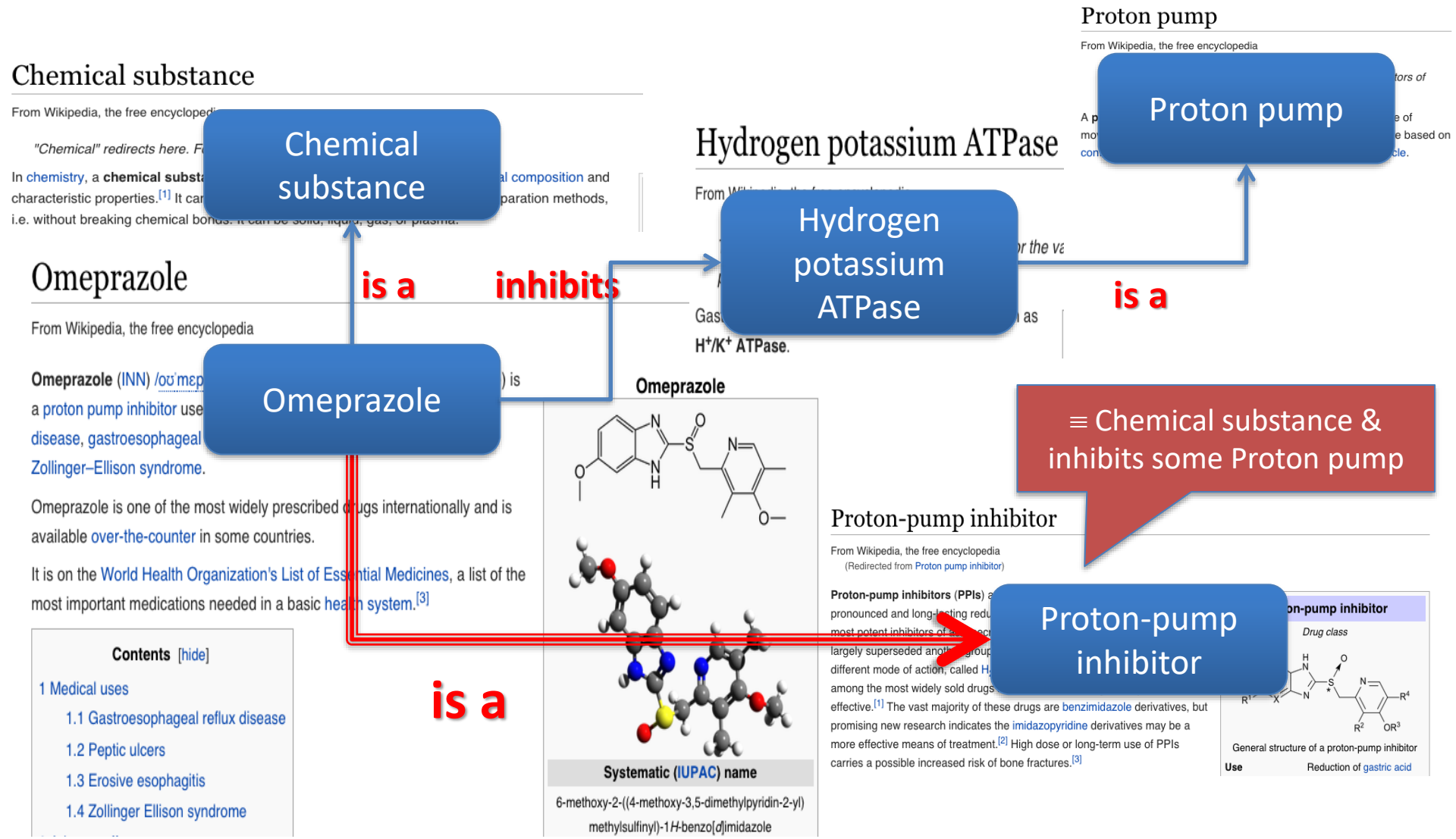


General structure of a proton-pump inhibitor

Use Reduction of gastric acid

Source: Brochhausen (2018)

Automatic inferences

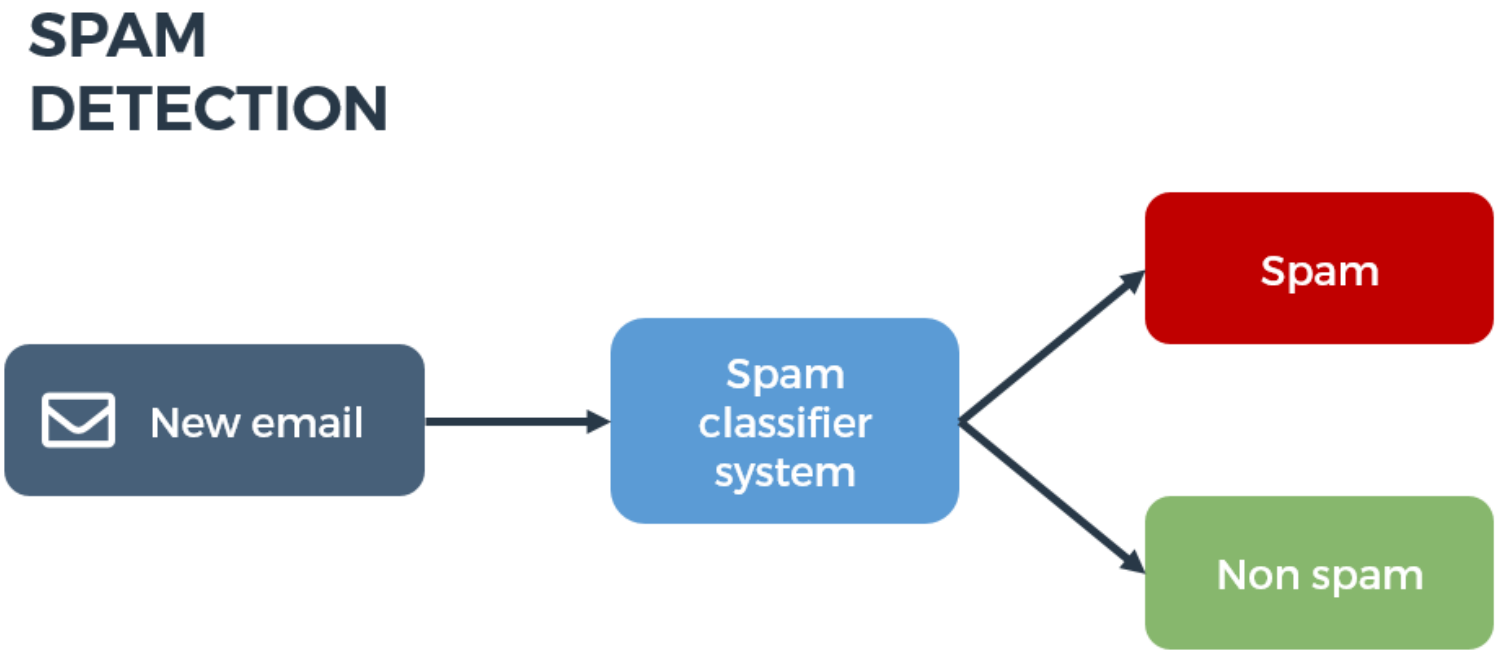


Brochhausen (2018)

II-1 Probability Inferences in Machine Learning

- It is a field with roots in 1950s
- Algorithms handle data to make decisions based on probabilities
- They learn from data sets, discover patterns, improve prediction
  - Supervised learning: there is an outcome variable to guide the learning
  - Unsupervised learning: describes how the data are clustered in patterns
- **Classification** is the supervised process of:
  - Determining the type of a variable
  - Producing an quantitative outcome

Probabilistic Inferences in Machine Learning



PART III- Discussion

Are claims about machine reasoning achievements really true?

OR

Are these claims a sort of hype?

Discussion

- Challenges faced by ML applications:
  - they are “data hungry”
  - they have limited capacity in transfer tests in slightly modified scenarios
  - they presume a largely stable world
  - they have no natural way to deal with hierarchical structures
  - they struggle with open-ended inferences
  - they are not sufficiently transparent, some call them “black-boxes”
  - they do not deal with background knowledge
  - they cannot distinguish causation from correlation
  - they work as an approximation, as any probabilistic model

Discussion

- Some requirements for ML applications work well:
  - A huge body of data must be available for training based on the best human performance
  - The input-output data must be similar, it requires patterns obtained in recurring processes (not erratic)
  - The data input must be abounding, millions of records are needed to represent the full variance

Discussion

- Limitations of ML applications regarding meaning:



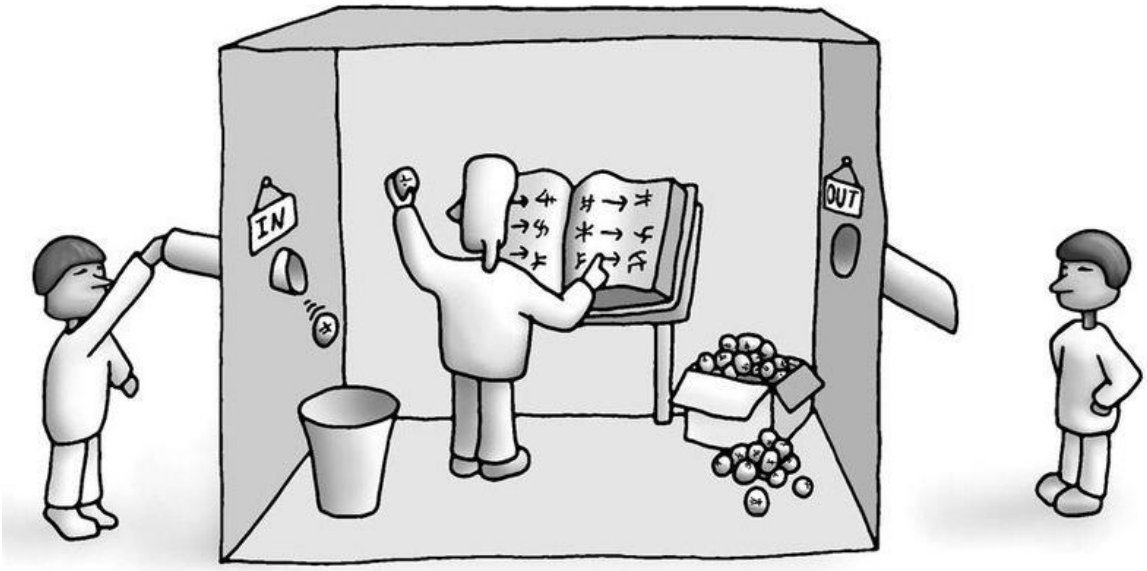
The ML models are agnostic
Three different objects recognized just as “blue things”

Source: Brochaussen (2018)

Discussion

- Requirements for ontology-based applications:
  - A large amount of time from experts
  - Axiomatized and populated ontologies are required for inferences
  - Much enginnering and expert´s work is required for maintenance
  - Much enginnering work is required for manual inspection
  - It need premises to generate a conclusion: only syllogistic reasoning

IV- Final remarks



Searle´s Chinese Room
from 1980s

Thank you!

mba@eci.ufmg.br

http://mba.eci.ufmg.br/

http://ppggoc.eci.ufmg.br/

V- References

Brochhausen, M. (2018). Lecture Notes. University of Arkansas For Medical Science. https://humanities.uams.edu/faculty/mathias-brochhausen-ph-d/

Ereshefsky, M. (2004). *The Poverty of the Linnaean Hierarchy*: A Philosophical Study of Biological Taxonomy. Cambridge University Press, Cambridge.

Genesereth, M.R., Nilsson, J. (1988). *Logical foundations of artificial intelligence*. Morgan Kaufmann , San Francisco.

Gorman, M., Sanford, J.J. (2004). Aristotelian Tradition. In Michael Gorman and Jonathan J. Sanford. (Eds.). *Categories*: Historical and Systematic Essays. The Catholic Univ of America Press, Washington D.C.

Gorman, M. (2004). Categories and Normativity. In Michael Gorman and Jonathan J. Sanford. (Eds.). *Categories*: Historical and Systematic Essays. The Catholic Univ of America Press, Washington D.C.

References

Hastie, T., Tibshirani, R., Friedman, J. (2008). *The Elements of Statistical Learning*: Data Mining, Inference, and Prediction. New York, Springer.

Hjørland, B. (2012). Is classification necessary after Google? Retrieved March 30, 2019 from http://www.iskoiberico.org/wp-content/uploads/2014/09/19-30_Hjorland.pdf.

Horridge et al. (2011). *The Cognitive Complexity of OWL Justifications*. Retrieved August 13, 2012, from http://ceur-ws.org/Vol-745/

Jansen, L. Classifications. In Katharine Munn, Barry Smith (Eds.). *Applied Ontology: an introduction,* Ontos-Verlag, Berlin.

Landgrebe, J., Smith, B. (2019). Making AI Meaningful Again. Retrieved March 2, 2019, from https://arxiv.org/abs/1901.02918

References

Russell, S., Norvig, P. (2017). *Artificial Intelligence: modern approach*. New York Prentice Hall

Oaksford, M., Chater, N. (2001). The probabilistic approach to human reasoning. *Trends in Cognitive Sciences*. Vol.5 No.8, 349-357.

Saracevic, T. (1996). *Interdisciplinary Nature of Information Science*. Retrieved August 30, 2017 from http://www.brapci.inf.br/_repositorio/2010/03/pdf_dd085d2c4b_0008887.pdf

Sattler, U., Stevens, R., Lord, P. (2014). *How Does a Reasoner Work*? Retrieved August 12, 2016 from http://ontogenesis.knowledgeblog.org/1486

References

Searle, J. (1980). Minds, Brains and Programs. *Behavioral and Brain Sciences*, Vol. 3, 417–57

Smith, B., Brogaard, B. (2003). Sixteen Days. *The Journal of Medicine and Philosophy*, Vol. 28, 45–78.