14.0 Semantic Web Reasoning

- Last week we saw ontologies as a powerful instrument for...
  - Representing knowledge
  - And reason about it!
- Ontologies, rules and logics form the middle layer of the proposed Semantic Web stack
  - Formal syntax
  - Formal semantics

While RDF/S (or at least the DLP bits) form a valid foundation for OWL, Datalog-style rule languages need other assumptions
  - Closed world semantics
  - Leads to full negation as failure (NAF)
  - ...
- Whereas DLP is only a subset of Horn rules
  - And if it is interpreted with Herbrand models and CWA, it is no longer suitable for OWL...
14.0 Semantic Web Reasoning

Hmmm… this leads to difficult questions…
- If you want to join the debate:
- Maybe rules on top of OWL…?!
14.1 The MeSH Ontology

- Qualifiers encode commonly used tags
  - Can be added to all other headings
  - e.g. viral, microbiol, epidemic, etc

- By using MeSH, concept maps can be visualized
  - Help to quickly assess a given topic

Visual dictionary uses co-occurrence of concepts in publications as weight indicator

- Typed links between concepts allow for "browsing"

- Also, can become easily very large and complex

MeSH is an example for enriched taxonomy manually modeled by domain experts

- Expert taxonomies are widely used, however, they come with problems
  - Inflexible and rigid structure representing just the authors view and knowledge
  - Hard to change once established, expensive to maintain
  - Hierarchical classification often not very practical
14.1 Hierarchical Expert Ontologies

• Example hierarchies:
  – Periodic Table of Elements, devised in 1869 by Dmitri Mendeleev
  – Probably the best classification scheme ever
  – But still, it is and was heavily disputed
    • Represented just the knowledge known by Mendeleev
    • e.g. initial version was missing noble gases
      – ...by the way, is Helium really a gas? It becomes solid when cooled...
    • Ordering scheme changed from weight to atomic number
    • Inserted and added rows / columns, added categories, etc
    • etc.

• Dewey Decimal Classification (DDC)
  – Proprietary system for library classification, developed by Melvin Dewey in 1876
  – Updated in varying intervals (currently 22nd revision)
  – Used by, e.g. Library of Congress
  – Organizes everything in 10 main classes, which are divided into 10 divisions, which have 10 sections
    • A less flexible variant of a system similar to the tree ID in MeSH
    • Strictly hierarchical

• One of the main problems in inflexibility and inability to further model relationships between entries
  – Also, all entries are considered to be co-equal
  – Until recently, classification for the top concept 200 – Religion looked like this:

• Currently, main categories are like

  000 – Computer science, information, and general works
  100 – Philosophy and psychology
  200 – Religion
  300 – Social sciences
  400 – Languages
  500 – Science and Mathematics
  600 – Technology and applied science
  700 – Arts and recreation
  800 – Literature
  900 – History and geography and biography

  – e.g. 025 is library management

• In the late 90ties, Yahoo! started to classify the World Wide Web
  – For this task, ontology experts where hired to create the classification hierarchy
  – Often, this classification was quite difficult and awkward
  – Also, links among entries were necessary between entries
    • Strict hierarchical modeling not sensible
14.1 Hierarchical Expert Ontologies

- Thus, the transition was made from **strictly hierarchical** to **linked hierarchical taxonomies**

14.1 Hierarchical Expert Ontologies

- In case of highly **unstructured domains**, capturing information in an hierarchical way becomes increasingly difficult
  - More and more links, hierarchy less and less useful
    - Idea: **Just omits the hierarchy** part and use only links
    - Folksonomies
    - Automatically generated Lightweight ontologies

14.1 Ontology Generation

- Of course the **manual creation** of ontologies is an expensive and error-prone process
  - Is there a possibility to **create ontologies automatically**?
  - It’s a current research question, but first approaches lead to semi-automatic procedures…
    - Basically all approaches mine statistical connections between terms…

14.1 Ontology Generation

- A major group for taxonomy creation are **natural language processing** approaches
  - Gathering simple typical phrases from full texts like “…such as…” or “…like e.g.,…” to find synonyms or subclasses
    - The surrounding noun phrases can be put into some (hierarchical) relationship
    - The belief in the correctness of derived classes and/or hierarchies can be supported by comparison to general ontologies like WordNet or counting co-occurrences e.g., in documents retrieved from Google

14.1 Ontology Generation

- Or, domain ontologies can be derived relying on simple statistics, e.g., **term co-occurrence**
  - Extract all **salient keywords** from each document
  - Keyword X **subsumes** keyword Y, if at least 80% of the documents in which Y occurs also contain X, and if X occurs in more documents than Y
  - Works only if a **sufficiently large** number of documents for a certain domain is given
• Please note, all these techniques are heuristics…
  – The Semantic Web does not really understand the contents of the pages (not yet!?)
  – But still, better than nothing…

  – Thus, the question arises: Can purely statistical approaches lead to reasonably intelligent results?

14.2 Wisdom of Crowds

• Just a little anecdote for the start:
  • Sir Francis Galton (1822–1911)
    – Victorian polymath with special interest in statistics
      • Established principles for correlation, deviation, and regression
      • Special interest in research methodologies of eugenics, heredity, genetics, and historiometry
        • Claim: Intelligence and leadership properties are inherited, and only few people possess them. And only those are able to lead and act intelligently.

14.2 Wisdom of Crowds

• In 1906, he visited a country fair which also featured an ox weighting betting contest
  – An ox is presented, everybody guesses how much the meat after slaughtering will weigh, closest bet wins
  – ~800 people participated
    • Farmers, housewives, castle experts, random visitors, children, etc
  – Galtons claim:
    • Experts will win, the other people will just guess nonsense, crowd consensus will be useless
  – Statistical analysis
    • Ox weighted 1,198 pounds, average guess of all people was 1,197 pounds, no single guess was better than crowd consensus
  – Galtons afterwards:
    • “The result seems more creditable to the trustworthiness of a democratic judgment than one might have expected.”

14.2 Wisdom of Crowds

• In 1968, the nuclear submarine USS Scorpion mysteriously disappeared
  – Search for the sub was hopeless and was abandoned
  – However, Dr. John Craven from Navy’s Special Projects continued the search with his team of mathematicians
  – Idea:
    • Provide all known evidence to a large group of peoples and teams
      • Submarine experts, salvage experts, oceanologists, mathematicians, ship captains, etc.
    • Each team should develop a theory to what happened and where the submarine was
      • Craven combined all theories (wildly diverse) using Bayes’s theorem
      • Submarine wreckage was immediately located 200 meters off the combined estimated location

14.2 Wisdom of Crowds

• This observation fueled a new research observing crowd decisions

  • Experiments and observed events
    – Bean guessing games
      • Crowd estimate always very good
    – “Who wants to be a millionaire” joker
      • 91% success rate vs. 65% expert success
    – Predicting outcomes of sport events
      • Aggregated bets are usually more accurate than any expert guess

14.2 Wisdom of Crowds

• Observation
  – Under certain restrictions large crowds of people are able to perform highly effective decisions
    • Far superior to nearly all singular decisions
    • Some care and control is required to prevent this approach from failing miserably
  – Further reading:
    • James Surowiecki: The Wisdom of the Crowds, 2004
14.2 Principles of Group Intelligence

- Group intelligence can effectively be used on three types of problems
- **Cognition Problems**
  - Judging and Processing Information
  - Examples
    - “Who will win Germany’s Next Top model?”
    - “How many beans are in the jar?”
    - “Which movie should one watch who liked Star Wars?”
    - “How can the music TOP-100 be classified into genres?”

- **Coordination Problems**
  - How to coordinate one’s own behavior with all others, knowing that they try to do the same?
  - Often, coordination problems encode cultural behavior
  - Examples
    - Navigating in heavy traffic
    - Using the seats in a lecture hall
    - How to figure out a good price for used items?

- **Cooperation Problems**
  - Get self-centered, distrustful people to work together for a greater good
  - Forming networks of trust without necessity of a central controller
  - Examples
    - The free market
    - Paying taxes
    - Dealing with pollution

- Diversity of opinion
  - Each person should have private information, even if it’s just an eccentric interpretation of the known facts
  - Opposing opinions usually increase the accuracy of group decisions by either…
    - Canceling each other's mistakes
    - Or fostering discussion among group members
  - However, it is important that everybody needs an understanding of the problem
    - You don’t need to ask kindergarten children about the potential cause of the SARS epidemic
    - Diversity means diversity of knowledgeable opinions, not any opinions!

- Groups being too homogeneous will not be able to tap into the power of their numbers
  - Too much of just the same comes up

- **Independence**
  - People’s opinions should not be determined by the opinions of those around them
  - The strength of group decision making comes from the diversity of opinion which will be lost, if the group members are not independent
  - Dominating members will affect the decisions of the other members
    - Hype bubbles
    - “Monkey see, monkey do”
14.2 Principles of Group Intelligence

- **Decentralization**
  - People are able to specialize and draw on local knowledge
  - Crucial to tap into people’s tacit knowledge
  - Specialization adds more diversity to the group
  - Specialist for a certain area provide more valuable input than non-specialists for a special problem
  - Using local knowledge allows for optimized solutions for special cases compared to central generic solutions

- **Problems with decentralization** without further adjustments
  - Wasted efforts
    - Many try to solve the same problem although it was already solved many times elsewhere
  - Crucial information does not propagate among the groups
    - Think of 09/11: most facts for predicting the incident were known, but scattered among all the intelligence agencies

14.3 Folksonomies

- **How can crowd intelligence be harnessed for our problems (i.e. dealing with knowledge in computer science)?**
  - Most popular example: Google PageRank
    - **Base idea:**
      - Each link from one page to another is a vote, i.e. the author thinks that the linked page is somewhat important
      - The more “votes” a page gets, the more important it is
      - Pages originating from important pages count more than those from unimportant ones
      - The votes thus propagate along all pages, encoding the common, aggregated belief of importance of all websites given by all website authors!
      - Incorporating the crowd knowledge given by page rank into traditional IR methods made Google the most successful search engine ever!

- **So, how to use crowds for actually modeling knowledge?**
  - **Observation around 2004:** People enthusiastically enjoyed tagging content on the web
  - Idea arose that these tags can be used to represent common, shared knowledge similar to ontologies
    - The folksonomy was invented!
      - Usually credited to Thomas Vander Wal
    - Also collaborative tagging, social classification, social indexing, and social tagging
14.3 Folksonomies

• What is tagging?
  – A tag is just some word which is assigned to some resource and represents some informal meta-data
  • Tags are usually freely chosen by the tagger

• The tags for a single resource can be represented by tag cloud
  – The bigger a tag appears, the more often it was used for this resource

• Now, a folksonomy could be build by, e.g. observing the co-occurrence of tags on resources

• What are folksonomies?
  – A folksonomy is a much weaker structure than description logic ontologies
  • No taxonomies and usually not even a vocabulary
  – A Folksonomy just link some tags to some other tags
    • The tags themselves as well as the links do not have to be necessarily meaningful
  – A Folksonomy represent the self-emergent semantics of the collaborative tagging effort

• Formal representation of folksonomies
  – A folksonomy T can be represented by a tripartite hypergraph $H(T) = <V, E>$
  • Vertices $V = A \cup C \cup I$ are partitioned into the disjoint sets
    – set $A$ of actors/users,
    – set $C$ of tags/concepts
    – set $I$ of instances/objects.
  • Each tag represents an edge between an actor, tag, and instance
    – $E = \{(a,c,i) \mid (a,c,i) \in T\}$

Ontologies are us: A unified model of social networks and semantics, Peter Mika, ISWC 2005
### 14.3 Folksonomies

- Based on the hypergraph of $T$, three **weighted bipartite graphs** can be generated
  - **Weight** represents how often the two diagrammed vertices of the bipartite graph had been connected by edges in the hypergraph
  - The graph $AC$ of actors and concepts
  - The graph $CI$ of concepts and instances
  - The graph $AI$ of actors and instances
  - E.g., see definition of AC:
    \[ AC = A \times C, E_{ac} = \{(a, c) \mid \exists i \in I (a, c, i) \in E\}, \]
    \[ w : E \rightarrow \mathbb{N}, \forall (a, c) \in E_{ac}, w(e) := |\{ i : (a, c, i) \in E\}| \]

### 14.3 Folksonomies

- The affiliation graphs can be folded into two lightweight ontologies
  - i.e. for the affiliation graph CI, we can get
    - The lightweight ontology of related concepts
    - The lightweight ontology of related instances
  - Those ontologies represent how strongly its contained entities are related
    - Similar to counting co-occurrence
    - Mathematically, this can be achieved by multiplying the matrix of the affiliation graph within its inverse, normalizing it with Jaccard-Coefficient, etc.

### 14.3 Folksonomies

- An excerpt from the delicious lightweight ontology graph

### 14.3 Folksonomies

- Resulting weighted graph CI
  - Also called affiliation graph
  - Optionally, a threshold can be applied to remove weak edges

### 14.3 Folksonomies

- Lightweight ontology for concepts in delicious

### 14.3 Folksonomies

- Some shortcomings
  - No controlled vocabulary
    - e.g., sciencefiction vs. Science Fiction vs. science_fiction
    - LOL, ROFL, knorpelfunky, etc.
  - Handling of synonyms and homonyms
    - IFS vs. Institute für Informations Systeme
    - Bachelor (degree) vs. Bachelor (unmarried male)
  - Questionable semantics of links
    - What does a link in a folksonomy mean? Does it mean something?
    - No formal reasoning possible!
14.3 Folksonomies

- delicious.com
  - Social bookmarking

14.4 The Web 3.0?

- Project 10X
  - "Industry Roadmap to Web 3.0 and Multibillion Dollar Market Opportunities"
    - Vast industrial report on semantic web business future
    - I.e. marketing blubber, but still realistic
  - Web 2.0: Connecting people
  - Web 3.0: Connecting knowledge
    - Add a "knowledge layer" on top of the internet
    - Finally realize the Semantic Web vision

- Claim: the support for creating the Web 3.0 is finally there
  - Semantic technologies embraced by many big players

14.4 The Web 3.0?

- Trends of Web 3.0
  - Semantic User Experience
    - "Intelligent user interfaces drive gains in user productivity & satisfaction"
  - Personalized, context aware, immersive human-computer interaction
  - Semantic Social Computing
    - "Collective knowledge systems become the next killer app"
  - Enrich Web 2.0 technologies (blogging, tagging, social networking, wikis, etc.) with semantic layers
    - Tag ontologies, semantic wikis, semantic blogs, etc

- Semantic Applications
  - "New capabilities, concepts of operation, & improved lifecycle economics"
  - Enhance enterprise-level off the shelf software (e.g. ERP, CRM, SCM, PLM, HR, etc) with knowledge layers
    - Ontology-driven discovery of documents
    - Policy-driven processes modeled using ontologies
    - Business logic modeling
    - Automated agents and advisors
  - Semantic Infrastructure
    - "Hardware for semantic software"
    - New immersive display technologies for better data interaction, specialized processors, mega-broadband internet, "everything connected"
  - Ubiquitous computing
14.4 The Web 3.0?

- The future internet in 2020? Web 4.0?

14.4 Web 3.0?

- Technologies for Web 3.

14 Thank You!

- I hope you enjoyed the lecture and learned at least some interesting stuff...
  
  Next semester's master courses:
  Multimedia Databases, XML Databases, GIS