Vorlesung

Informationswissenschaft und Informationssysteme

Hans Uszkoreit & Brigitte Jörg

Ontologies

- What is an ontology
- Examples
- Digging Deeper
- Outlook

by Hans-Ulrich Krieger

What is an ontology ?

Ontology [Greek]: most fundamental branch of general meta-physics, dealing with the study of existence (science of being; Aristotle, 384BC–322BC)

first occurrence of the term ontologia as we use it today by Jacob Lorhard (1561–1609; Jacobo Lorhardo, Jacobus Lorhardus) in first edition of Ogdoas Scholastica (1606)

discipline can be subdivided into

- formal ontology (or universal science)
- material ontology

Formal Ontology

question: what are the truth-determining foundations of general metaphysics, i.e., what are the most general rules directing our decisions, leading to more specialized rules (e.g., in medicine):
first principles

- Law of Identity
  A = A: an axiom in most logics

- Law of Excluded Middle
  either P or ¬P

- Law of Non-Contradiction
  proof by contradiction: $(\neg P \Rightarrow (R \land \neg R)) \Rightarrow P$

by Hans-Ulrich Krieger

Material Ontology

what are the fundamental categories of being? (Aristotle)

more general view: find out what entities and what types of entities exist!

similar to the idea of first principles: start with Being (does not need any definition), and add subcategories, such as Substance

what does it mean for an entity to be member of a certain category? sharing prototypical values for category-specific properties!

by Hans-Ulrich Krieger

Reappearance of the Wheel

Aristotle's theory of categories and classification "reappears" in philosophy and many other scientific disciplines:

- biology
- ...
- CL, AI, CS, LT, ...
  - (computational) linguistics
  - artificial intelligence
  - computer science
  - information science, lexicography, semantic web, ...

by Hans-Ulrich Krieger

What is an ontology: Tom Gruber (1993)

A conceptualization is an abstract , simplified view of the world that we wish to represent for some purpose. . . . An **ontology** is *an explicit specification of a conceptualization*.

. . . When the knowledge of a domain is represented in a *declarative formalism*, the set of objects that can be represented is called the universe of discourse. This set of *objects , and the describable relationships* among them, are reflected in the *representational vocabulary* with which a knowledge-based program represents knowledge.

by Hans-Ulrich Krieger

What is an ontology: Tom Gruber (1993)

an ontology is a description of objects (categories & individuals) and relationships between objects

1+is-a relation: *taxonomy*; 1+2: *thesaurus*

1. categories/concepts/classes/types: `Man`
2. (built-in) relations between categories: `Man subclassOf Human`
3. individuals/instances: `peter, mary`
4. relations/roles between individuals: `peter isMarriedTo mary`

What is missing here? **semantics!** (later)

by Hans-Ulrich Krieger

Why are we interested in Ontologies ?

- pure **epistemological** aspects—no practical interest in running systems
  - build models of (specific parts of) the world
  - find encoding that conforms with taken observations
  - good model should predict facts not encountered so far
  - questions:
    - what can be encoded in the representational vocabulary and what can not?
    - what is the computational complexity of the representation language?
    - is the language decidable?
- very **practical** aspects → next slide

by Hans-Ulrich Krieger

Application Areas

- query expansion in IR & QA
- DB access & ontology retrieval
- word sense disambiguation
- ontology population through IE
- language-specific inferences on lexical semantic representation
- general inferences dealing with world knowledge

by Hans-Ulrich Krieger

Examples

- Thesauri
- WordNet
- FrameNet
- SUMO / MILO
- Description Logics & OWL

Merriam-Webster Online Thesaurus

Word: human

Function: adjective

Text: relating to or characteristic of human beings (it's human nature to care about what people think of us)

Synonyms: mortal, natural

Related Words: anthropoid, hominid, humanlike, humanoid

Near Antonyms: angelic (or angelical), divine, godlike, superhuman, supernatural; immortal, omnipotent, omniscient; animal, beastly, bestial, brute; inhuman, robotic

Antonyms: nonhuman

Merriam-Webster Online Thesaurus, cont.

Word: human

Function: noun

Text: a member of the human race (humans are the only mammals not endowed with a natural defense against the elements, such as fur or a thick hide)

Synonyms: being, bird, body, creature, customer, devil, guy, head, individual, life, man, mortal, party, person, scout, sort, soul, specimen, thing, wight

Related Words: hominid, homo, humanoid; brother, fellow, fellowman, neighbor; celebrity, personage, personality, self, somebody

Near Antonyms: animal, beast, brute

by Hans-Ulrich Krieger

WordNet – Hypernyms of Human

WN *hierarchically* organizes nouns, verbs, adjectives, and adverbs into synonym sets which refer to lexical concepts (155,327 unique strings & 117,597 synsets in WordNet 3.0)

Sense 1 / noun: *a human being*

person, individual, someone, somebody, mortal, human, soul
    => organism, being
        => living thing, animate thing
            => object, physical object
                => entity
            => causal agent, cause, causal agency
                => entity

by Hans-Ulrich Krieger

WordNet – Hypernyms of Human, cont.

Sense 2/noun: *any living or extinct member of the family Hominidae*

homo, man, human being, human

=> hominid

=> primate

=> placental, placental mammal, eutherian, eutherian mammal

=> mammal

=> vertebrate, craniate

=> chordate

=> animal, animate being, beast, brute, ...

=> organism, being

=> living thing, animate thing

=> object, physical object

=> entity

Relations we are interested in w.r.t. concept C

- **synonyms** concepts having the same meaning as C
- **antonyms** concepts that do not share any properties with C
- **hypernyms** concepts that are more general than C
- **hyponyms** concepts that are more specific than C
- **holonyms** concepts that contain C as a part
- **meronyms** concepts that are part of C

FrameNet – Human, again

FN lists semantic and syntactic combinatory possibilities (valences) of each word in each of its senses (> 10,000 lexical units; ≈ 800 hierarchical semantic frames)

two lexical units for *human*: `human being.n` and `human.n`

but semantic frame is `People`

several "subclasses" of `People`, e.g., `People_by_age`

binary relations, connecting frames: `Inherits_From, Uses, ...`

example: `People_by_age Inherits_From People` ("specialization")
`People_by_age Uses Age` ("properties")

SUMO & MILO

Suggested Upper Merged Ontology: very basic concepts & axioms
(similar upper ontologies: DOLCE, PROTON)

higher-order LISPish specification language SUO-KIF

```
(instance instance BinaryPredicate)
(subrelation immediateInstance instance)
(instance immediateInstance AsymmetricRelation)

(=> (immediateInstance ?ENTITY ?CLASS)
     (not (exists (?SUBCLASS)
          (and (subclass ?SUBCLASS ?CLASS)
               (not (equal ?SUBCLASS ?CLASS))
               (instance ?ENTITY ?SUBCLASS)))))
```


by Hans-Ulrich Krieger

SUMO & MILO, cont.

MId-Level Ontology: bridges between the abstract content of SUMO and various domain ontologies

all ontologies together: 20,000 terms and 60,000 axioms
partial inference support via Vampire

```
(subclass HumanSlave Human)

(=> (instance ?SLAVE HumanSlave)
     (exists (?PERSON)
         (and (instance ?PERSON Human)
           (not (equal ?PERSON ?SLAVE))
           (possesses ?PERSON ?SLAVE))))
```

SUMO & MILO – That Human Thing, Again

mappings of concepts to WordNet lexicon
example *human*: found the two senses from WordNet

```
(partition Human Man Woman)
(subclass Human CognitiveAgent)
(subclass Human Hominid)
(subclass Man Human)

(<=>    (attribute ?PERSON Unemployed)
        (and (instance ?PERSON Human)
             (forall (?ORG)
              (not (employs ?ORG ?PERSON)))))
```

Description Logics

family of logic-based knowledge representation formalisms DL
example: OWL (later!)

descendants of semantic networks and KL-ONE

describe domain in terms of concepts, roles, and individuals

complex expressions through concept-forming constructors

$$\texttt{HumanSlave} \equiv \texttt{Human} \sqcap \exists\ \texttt{possesses}^{-1}.\ (\texttt{Human} \sqcap \neg\ \texttt{Slave})$$

$$\texttt{HappyFather} \equiv \texttt{Man} \sqcap \forall\ \texttt{hasChild}.\ (\texttt{Doctor} \sqcap \exists\ \texttt{hasFriend}.\ (\texttt{Rich} \sqcup \texttt{Famous}))$$

by Hans-Ulrich Krieger

Description Logics, cont.

model-theoretic semantics (decidable 2-var fragment of FOL)

sound & complete decision procedures

highly optimized implemented systems

increasing importance for

- Tim Berners-Lee's vision of a Semantic Web
- language technology (ontology-based information systems)
- artificial intelligence (multi-agent systems, user modeling)
- computer science (deductive, object-oriented data bases)

by Hans-Ulrich Krieger

Recap: What is an Ontology

similarities between examples indicate that

- I take a liberal stance here what an ontology is
- we always construct ontologies when conceptualizing a domain
  1. categories / concepts / classes / types
  2. distinguished sub / super relationship
  3. individuals / instances / entities
  4. relations / roles / properties / attributes

- but: formal ontology languages must address
  – semantics: well-defined (yes)
  – decidability: sound (yes) & complete calculus (yes .. no)
  – tractability: average-case problems (yes .. no)

by Hans-Ulrich Krieger

The Semantic Web Vision

(syntactic) Web made possible through established standards: TCP/IP, HTTP, HTML, ...

1st generation: mostly handwritten HTML pages

2nd generation: very often machine-generated active pages

next generation (we're just here!): resources should be more accessible to automated processes

- to be achieved via semantic markup
- metadata annotations, describing content/function

coincides with Tim Berners-Lee's vision of a Semantic Web

Semantic Web and Ontology

semantic markup must be meaningful to automated processes

ontologies will play a key role here

- source of precisely defined terms (vocabulary)
- can be shared across applications and humans

increased formality facilitates machine understanding

very important: standards!
long road:
XML, URI, **RDF**, **RDFS**, DAML & OIL, **OWL**, **SWRL**, ...

by Hans-Ulrich Krieger

RDF: Resource Description Framework

- general-purpose language for representing information
- provides a lightweight ontology system
- enabling technology for the Semantic Web
- XML exchange syntax (but also N3, N-Triples)
- RDF data model: triple
- idea: everything can be represented as a triple

by Hans-Ulrich Krieger

RDF, cont.

- **triple**: ⟨ subject, predicate, object ⟩
- subject, predicate, object: URIs or XSD literals (or again triples: reification)
- **URI**: Uniform Resource Identifier (≈ Web identifier) e.g., http://www.w3.org/2002/07/owl#intersectionOf
- **XSD**: XML Schema Datatypes typed **literals**, e.g., "2.4"^^xsd:decimal

RDFS: RDF Schema

- describes how to use RDF to describe RDF vocabularies
- defines other built-in RDF vocabulary (domain, subClassOf)
- class & property system similar to OOPL (e.g., Java)
- RDF(S) semantics via **axiomatic triples & entailment rules** (Hayes 2004), e.g.,
  - $\langle$rdf:type, rdf:type, rdf:Property$\rangle$
    $\langle$rdfs:subPropertyOf, rdfs:subPropertyOf, rdfs:subPropertyOf$\rangle$
  - $\langle$?p, rdfs:domain, ?d$\rangle$ $\wedge$ $\langle$?s, ?p, ?o$\rangle$ $\Rightarrow$ $\langle$?s, rdf:type, ?d$\rangle$
    $\langle$?i, rdf:type, ?d$\rangle$ $\wedge$ $\langle$?d, rdfs:subClassOf, ?c$\rangle$ $\Rightarrow$ $\langle$?i, rdf:type, ?c$\rangle$

by Hans-Ulrich Krieger

OWL

decidable instance of the description logics family (FOL fragment)

well-founded set-theoretical semantics

outcome of the DAML+OIL W3C standardization

de facto standard today to specify ontologies

RDFS-based syntax and ontological primitives
e.g., `rdfs:subClassOf`

fine-grained, more complex means as in RDFS
e.g., `owl:intersectionOf`

uses XML/RDF exchange syntax

ontology is a set of axioms describing classes and properties

OWL

Three increaingly expressive sublanguages

- base $\mathit{ALC}_R{}^+ = S$
- **OWL Lite:** sound & complete, decidable
  reasoning services: ExpTime (worst case)
  optimized implementations: tableaux algorithms
- **OWL DL:** sound & complete, decidable (NExpTime)
  extends OWL Lite with disjunction & negation, cardinality constraints, and nominals
- **OWL Full:** reasoning usually undecidable

Class vs. Instance

classes & class properties (KL-ONE: **TBox**; DB: Schema)

```
owl:Class
  owl:equivalentClass rdfs:subClassOf
  owl:intersectionOf owl:unionOf owl:complementOf
  owl:disjointWith
owl:ObjectPoperty owl:DatatypeProperty
  rdfs:subPropertyOf owl:equivalentProperty
  rdfs:domain rdfs:range
```

instances or individuals (**ABox**; DB: complete knowledge)

`owl:sameAs owl:differentFrom owl:AllDifferent`

plus instantiated object/datatype properties

plus: `rdf:type`, ...

ontology = TBox + ABox (+ RBox)

Reference

… mehr bei

-> Hans-Ulrich Krieger

Links and Books

Tom Gruber's article: www.-ksl.stanford.edu/kst/what-is-an-ontology.html

RDF & OWL recommendations of W3C: www.w3.org/2004/01/sws-pressrelease

Resource Description Framework: www.w3.org/RDF/

RDF Schema: www.w3.org/TR/rdf-schema/

OWL: www.w3.org/2004/OWL/

WordNet: wordnet.princeton.edu/

FrameNet: framenet.icsi.berkeley.edu/

SUMO & MILO ontology: www.ontologyportal.org/

PROTON ontology: proton.semanticweb.org

by Hans-Ulrich Krieger

Links and Books

DOLCE ontology: dolce.semanticweb.org

Protégé: protege.stanford.edu/

OWLIM: www.ontotext.com/owlim

SWRL: www.w3.org/Submission/SWRL/

Ontology resources: www-ksl.stanford.edu/kst/ontology-sources.html

more resources: protege.cim3.net/cgi-bin/wiki.pl?ProtegeOntologiesLibrary

OWL-Time: www.w3.org/TR/owl-time/

FaCT: www.cs.man.ac.uk/~horrocks/FaCT/

RACER: www.sts.tu-harburg.de/%7Er.f.moeller/racer/

Pellet: http://clarkparsia.com/pellet/

Knowledge Interchange Format: logic.stanford.edu/kif/

by Hans-Ulrich Krieger

Links & Books

Cyc: www.cyc.com/

Description Logics homepage: http://www.dl.kr.org/

W3C group Semantic Web: www.w3.org/2001/sw/

Web Ontology Working Group: www.w3.org/2001/sw/WebOnt/

T. Berners-Lee et al.: The Semantic Web, Scientific American. www.sciam.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21 F. Baader et al.:

Description Logic Handbook, Cambridge University Press; see also www.inf.unibz.it/~franconi/dl/course/.

P. Hayes: RDF Semantics, 2004 (http://www.w3.org/TR/rdf-mt/).

H. ter Horst: Combining RDF and Part of OWL with Rules: Semantics, Decidability, Complexity. ISWC 2005, 668–684.

J.W. Lloyd: Foundations of Logic Programming, Springer.

Links & Books

M. Huth & M. Ryan: Logic in Computer Science, Cambridge University Press.

M. Tarnowski: Mathematische Grundlagen der formalen Linguistik, IWBS Report 174, IBM.

B.H. Partee et al.: Mathematical Methods in Linguistics, Kluwer.

G. Smolka: Logische Programmierung. www.ps.uni-sb.de/courses/lp-course93.html