kc-viz: A Novel Approach to Visualizing and Navigating Ontologies

Enrico Motta[1], Paul Mulholland[1], Silvio Peroni[2], Mathieu d'Aquin[1], Jose Manuel Gomez-Perez[3], Victor Mendez[3], Fouad Zablith[1]

[1]The Open University, [2]University of Bologna, [3]iSOCO

TopBraid – Sumo/SUMO.owl – TopBraid Composer ME – /Users/EnricoMotta/Desktop/Topbraid Workspace

Classes: rdfs:Resource (0 + 1564); Entity (0 + 541); Abstract (0 + 541); Attribute (0 + 52); Proposition; Quantity (0 + 41); Relation (0 + 447); SetOrClass (0 + 1); Physical; InheritableRelation (24); IntentionalRelation (1 + 6); ObjectAttitude (2); PropositionalAttitude (4); owl:AllDifferent; owl:Thing; rdf:Statement; rdfs:Class (657 + 46)

Navigator: SUMO.owl; SUMO-Large; .project; SUMO.owl; Sweet; swrc; T1; test12; Test22; .project

Transitway, Region, Food, SelfConnectedObject, CorpuscularObject, Agent, Substance, Collection

SUMO (file:/Applications/NeOn%20Toolkit/NeOnToolkit-2.3.1-B264-macosx/Nc

Active Ontology | Entities | Classes | Object Properties | Data Properties | Individuals | OWLViz | DL Query

Class hierarchy: Object

owl:Thing
- Entity
  - Abstract
    - Attribute
    - Proposition
      - Argument
      - FieldOfStudy
      - Procedure
    - Quantity
    - Relation
      - BinaryRelation
      - List
      - PartialValuedRelation
      - Predicate
      - ProbabilityRelation
      - QuaternaryRelation
      - QuintaryRelation
        - QuaternaryFunction
        - QuintaryPredicate
      - RelationExtendedToQuantities
      - SingleValuedRelation
      - SpatialRelation
      - TemporalRelation
      - TernaryRelation
        - BinaryFunction
        - TernaryPredicate
      - TotalValuedRelation
      - VariableArityRelation
  - SetOrClass
  - Physical

OWLViz: Object

Asserted model | Inferred model

Entity, Physical, Object, Agent, Collection, Region, SelfConnectedObject, CommercialAgent, SentientAgent, Organism, Group, GeopoliticalArea, AstronomicalBody, GeographicArea, Hole, Transitway, Food, Substance, CorpuscularObject

NeOn Toolkit

Ontology Navigator: food_products.owl; food_taxonomy.owl; free-base.owl; generic.owl; philips-test.owl; philips-test2.owl; Classes; Thing; Abstract; Agent; SpatialThing; EntityInSpace; Location; AbsoluteLocat; AbstractAb; Absolute; Abstract; Geograp; LineSegr; Geographic; PhysicalLoc; RelativeLocatic; TemporalThing; Assembly

temporalExtentEnds; hasTimeStamp; temporalExtentStarts

Entity Properties | KC-Viz | Ontology Visualizer

Zoom: 100% Rotate: 0° Search History Legend Tooltips

Event, SituationalContext, AssemblyComponent, Assembly, TemporalThing, Measurement, Product, TemporalEntity, EntityInSpace

Legend: Class

Navigation history: philips-test2.owl → TemporalThing

Flying over file:/Applications/NeOn%20Toolkit/NeOnToolkit-2.3.1-B264-macosx/NeOn%20Toolkit.app/Contents/MacOS/workspace/SUMO/SUMO.owl

SUMO.owl

Focus

Class List | Concise Format

IntentionalProcess (15)
- SocialInteraction
- IntentionalPsychologicalProcess
- OrganizationalProcess
- ContentDevelopment
- Making
- Guiding
- RecreationOrExercise
- Repairing
- Maneuver
- Poking
- Keeping
- Pursuing
- Maintaining
- Looking
- Listening

Search

0.08077830188679246

Zoom + −

Issues with current interfaces for ontology engineering

- Poor ability to provide ‘gestalt’ views of an ontology
  - Essential to support sensemaking, esp. when exploring ontologies created by other people
  - This issue is related to the lack of good support for **abstraction** and **saliency**
- Poor support for ontology navigation
  - Entity browser OK but limited
  - Graphical interfaces rather inflexible in their support for exploration
    - E.g., either expand only direct subclasses or all subtree
  - Some very interesting ideas (e.g., CropCircles), but limited in functionalities
- Observational studies (Dzbor et al., 2006) indicate low levels of user satisfaction with interface support in ontology engineering tools
  - Esp. lack of support for selective visualization of ontology parts, summaries, and overviews [4].
  - Problems affect especially naïve users.

KC-Viz: Key concept based visualization

- Uses an ontology summarization algorithm (KCE) as a way to provide ‘conceptual’ abstraction facilities
  - In contrast with all other tools which try to support abstraction through graphical means (if they support abstraction at all)
  - KCE parameters can be customised by the users
- Rich set of options for exploring/hiding parts of an ontology
- Scales up to large ontologies
- Support for graphical layout customization, history browsing, loading/saving snapshots, etc..
- Integration with other components of NeOn Toolkit
  - Esp Entity Browser and Inspector

Automatically identifying Key Concepts

ABSTRACT-INFORMATION ACADEMIC-ADMIN-STAFF-MEMBER ACADEMIC-DEGREE ACADEMIC-STAFF-MEMBER ACADEMIC-SUPPORT-STAFF-MEMBER
ACADEMIC-SUPPORT-UNIT ACADEMIC-UNIT ACTIVITY ACTIVITY-STATUS ADDRESS ADMINISTRATIVE-STAFF AFFILIATED-PERSON AGENT-TECHNOLOGY AMOUNT-OF-MONEY
APPELLATION APRIL ARTICLE-IN-A-JOURNAL ARTICLE-REFERENCE ASSISTED-PERSON ATTENDING-A-CONFERENCE ATTENDING-AN-EVENT AUGUST AWARD AWARDING-
BODY BINARY-RELATION BOOK BOOK-PUBLISHING BOOK-REFERENCE BOOK-SECTION-CONTRIBUTION BOOK-SECTION-CONTRIBUTION-REFERENCE
BUSINESS-AREA CALENDAR-DATE CAPITAL-CITY CHARITABLE-ORGANIZATION CITY CIVIL-SERVICE
CLASS CLASS-PARTITION COMPANY COMPOSITE-PUBLICATION COMPOSITE-PUBLICATION-REFERENCE COMPUTING-TECHNOLOGY
CONFERENCE CONFERENCE-PROCEEDINGS CONFERENCE-PROCEEDINGS-REFERENCE CONFERRING-AN-AWARD COORDINATE-LOCATION COUNTRY
CURRENCY DAILY-NEWSPAPER DAY DAY-IN-TIME DECEMBER DEGREE DISTANCE-TEACHING-UNIVERSITY DURATION EDITED-BOOK EDITED-BOOK-REFERENCE EDITOR
EDUCATIONAL-EMPLOYEE EDUCATIONAL-ORGANIZATION EDUCATIONAL-ORGANIZATION-UNIT ELECTRONIC-PUBLICATION EMAIL-ADDRESS EMAIL-MESSAGE EMPLOYEE
EMPLOYMENT-CONTRACT-TYPE ENUMERATED-SET EVENT EVENT-INVOLVING-MOVEMENT EVENT-INVOLVING-PRODUCTION FEBRUARY
FEBRUARY-IN-LEAP-YEARS FINANCIAL-AWARD FREELANCE-SOFTWARE-DEVELOPER FUNCTION GENDER GENERALISED-MEANS-OF-TRANSPORT
GENERALIZED-TRANSFER GENERIC-AGENT GENERIC-AREA-OF-INTEREST GEOGRAPHICAL-DATUM GEOGRAPHICAL-REGION GEOPOLITICAL-ENTITY
GIVING-A-TALK GOVERNMENT GOVERNMENT-ORGANIZATION GRANT GRAPHIC-DESIGNER HARDWARE-PLATFORM
HARDWARE-TECHNOLOGY HIGHER-EDUCATIONAL-ORGANIZATION HIGHER-EDUCATIONAL-ORGANIZATION-EMPLOYEE HOUR-IN-TIME IMPLEMENTED-SYSTEM INDIVIDUAL
INDUSTRIAL-ORGANIZATION INFORMATION-BEARING-OBJECT INFORMATION-EXTRACTION-TECHNOLOGY INFORMATION-TRANSFER-EVENT INFORMATION-TRANSFER-
MEDIUM INTANGIBLE-THING INTEGER ISBN-NUMBER ITEM-IN-A-COMPOSITE-PUBLICATION JANUARY JOURNAL JOURNAL-ISSUE JOURNAL-ISSUE-REFERENCE JULY JUNE
KNOWLEDGE-MODELLING-TECHNOLOGY KNOWLEDGE-PROGRAMMING-LANGUAGE LANGUAGE-ENGINEERING-TECHNOLOGY LEARNING-CENTRED-ORGANIZATION
LECTURER-IN-ACADEMIA LEGAL-AGENT LETTER LIST LOCAL-DISTRICT LOCATION MAGAZINE MARCH MAY MEETING-TAKING-PLACE MESSAGE METHOD MINUTE-IN-TIME
MONTH MONTH-IN-TIME MULTIMEDIA-DESIGNER MULTIMEDIA-TECHNOLOGY MUNICIPAL-UNIT NEGATIVE-INTEGER NEGATIVE-NUMBER NEWS-ITEM NEWSLETTER
NEWSPAPER NON-NEGATIVE-INTEGER NON-NEGATIVE-NUMBER NON-PROFIT-ORGANIZATION NOVEMBER NUMBER OCTOBER
ONTOLOGY-SPECIFICATION-LANGUAGE OPERATING-SYSTEM OPERATIONAL-ONTOLOGY-SPECIFICATION-LANGUAGE ORGANIZATION ORGANIZATION-SIZE ORGANIZATION-
UNIT PAPER-IN-PROCEEDINGS PAPER-IN-PROCEEDINGS-REFERENCE PARTNERSHIP PERIODICAL-PUBLICATION PERSON PHD-STUDENT
PHYSICAL-QUANTITY POLITICAL-ORGANIZATION POSITIVE-INTEGER POSITIVE-NUMBER POSTAL-ADDRESS PRIVATE-COMPANY
PROCEEDINGS PROCEEDINGS-REFERENCE PROFESSOR-IN-ACADEMIA PROFIT-ORGANIZATION PROGRAMMING-ENVIRONMENT PROGRAMMING-LANGUAGE
PROJECT PROJECT-OFFICER-IN-ACADEMIA PROJECT-OUTPUT PUBLIC-COMPANY PUBLICATION PUBLICATION-REFERENCE
PUBLICATION-TYPE-EVENT PUBLISHING-HOUSE QUANTITY R-AND-D-INSTITUTE R-AND-D-INSTITUTE-WITHIN-EDUCATIONAL-ORGANIZATION R-AND-D-INSTITUTE-WITHIN-
LARGER-ORGANIZATION READER-IN-ACADEMIA REAL-NUMBER RECORDED-AUDIO RECORDED-VIDEO REFERENCE-TO-ITEM-IN-A-COMPOSITE-PUBLICATION RELATION
RESEARCH-AREA RESEARCH-ASSISTANT-IN-ACADEMIA RESEARCH-FELLOW-IN-ACADEMIA RESEARCH-STAFF-MEMBER RESEARCHER SCHOOL SCRIPTING-LANGUAGE
SECOND-IN-TIME SECRETARY SELF-EMPLOYED-PERSON SENDING-AN-EMAIL SENIOR-LECTURER-IN-ACADEMIA SENIOR-RESEARCH-FELLOW-IN-ACADEMIA SEPTEMBER
SERIAL-PUBLICATION SERVER-TECHNOLOGY SET SMALL-OR-MEDIUM-SIZED-ORGANIZATION SOCIAL-GATHERING SOFTWARE-DEVELOPER SOFTWARE-STATUS SOFTWARE-
TECHNOLOGY SOFTWARE-VISUALIZATION-TECHNOLOGY SPECIFICATION-LANGUAGE SPECIFICATION-OR-COMPUTING-LANGUAGE STRING STUDENT SYSTEM-
ADMINISTRATOR TANGIBLE-THING TECHNICAL-REPORT-REFERENCE TECHNOLOGY TEMPORAL-THING THESIS-REFERENCE THING TIME-INTERVAL TIME-MEASURE TIME-
POINT TIME-POSITION TIMEZONE TOWN TRANSPORTATION-DEVICE UNARY-RELATION UNIT-OF-MEASURE UNIVERSITY UNIVERSITY-FACULTY URL VILLAGE VISITING-
RESEARCHER WEB-BASED-EDITOR WEB-BASED-SYSTEM WEB-BROWSER WEB-REFERENCE WEB-SITE WEB-TECHNOLOGY WEEK WORK-STATUS WORKING-PERSON WORKSHOP
WORKSHOP-PROCEEDINGS WORKSHOP-PROCEEDINGS-REFERENCE YEAR YEAR-IN-TIME

→

Project
Researcher
Phd Student
Technology
Publication
Educational Organization
.
.

Evidence that human experts are able to summarize ontologies very effectively

KNOWLEDGE MEDIA
KMi
INSTITUTE

Identifying key concepts: Approach

- Integration of cognitive criteria with lexical statistics, formal and topological criteria
  - Natural categories (Rosch, 1978)
    - information rich concepts that are ‘basic’ from a cognitive standpoint
    - E.g., dog, cat, chair, etc..
  - Density
    - information rich concepts from a formal standpoint
    - i.e., concepts rich in attributes, instances, or subclasses,
    - We use both local and global density measures
  - *Popularity*
    - *Lexical statistics*
    - *Familiar words tend to be more descriptive than unfamiliar one*
    - *We use both global and local popularity measures*
  - Best ontology coverage (topological)
    - We want to ensure that for each concept C in the ontology, there is a key concept Ki, such that either $C \subseteq Ki$ or $Ki \subseteq C$

Empirical evaluation of KCE

- 4 Ontologies – 8 Experts
- Experts showed on average a **74.68%** agreement ratio
- KCE showed a **72.08%** agreement ratio with the experts
  - Practically KCE could not be distinguished from experts in its ability to extract the best descriptors of an ontology

Peroni, S., Motta, E., and d'Aquin, M. (2008). Identifying key concepts in an ontology through the integration of cognitive principles with statistical and topological measures. *3rd Asian Semantic Web Conference*, Bangkok, Thailand, 8-11 December, 2008. **Best Paper Award**

Demo

Evaluation

Evaluation: Basics

- 21 subjects, 3 groups
  - Neon Toolkit without KC-Viz; Protégé/OwlViz; NeOn Toolkit with KC-Viz
- Process
  - Subjects filled a questionnaire, answering questions about their expertise in ontology engineering, KR languages, and ontology engineering tools
    - E.g., NeOn, Protégé, etc..
  - Tutorial (using pizza ontology)
  - 1 warm up task (using AKTLite Ontology)
  - 4 tasks (15 min per task) with SUMO-Lite Ontology (630 classes)
    - http://www.ontologyportal.org/translations/SUMO.owl.
    - Upper-level ontology with generic concepts
      - Reasonable to expect subjects to be familiar with the contents
      - Right level of size/complexity for the task
  - SUS Usability Questionnaire
  - KC-Viz demo (for those who did not use KC-Viz for the test)

Tasks

- **T1.** Which class has the highest number of direct subclasses in the ontology?
- **T2.** What is the most developed (i.e., has the biggest subtree) subclass of class Quantity found in the ontology at a concrete level of granularity (i.e., do not consider abstract classes which have the term ‘quantity’ in their id)?
- **T3.** Find three subclasses of Agent, at the most abstract level possible (under Agent of course), which are situated at the same level in the hierarchy as each other, and are also subclasses of CorpuscularObject.
- **T4.** We have two individual entities (a particular copy of the book War&Peace and a particular 5p coin). Find the most specific classes in the ontology, to which they belong, say P1 and P2, and then identify the most specific class in the ontology, say C1, which is a superclass of both P1 and P2 – i.e., the lowest common superclass of both P1 and P2.

Rationale for task specification

- No benchmark for this kind of experiment
- Reproducibility criterion requires reasonably fine-grained tasks with clear performance criteria
- Need to avoid task specification which mapped directly to KC-Viz features
  - E.g., “find most important concepts”
- Navigation is an essential element of a sensemaking process
- Some degree of intersection with tasks used in previous studies – e.g., (Parsia, 2005)

Results

	NTK		OWLViz		KCViz	
	mean	s.d.	mean	s.d.	mean	s.d.
T1	12:03	02:51	12:19	04:16	05:10	03:07
T2	06:43	04:45	07:20	03:55	04:03	02:15
T3	11:00	05:31	07:24	04:27	06:25	05:06
T4	08:01	05:56	08:23	05:28	06:17	05:15
Total	37:47	15:02	35:26	15:36	21:55	10:32

- KC-Viz subjects faster on all four tasks
- However difference significant only on Task 1
  - although approaching significance also on Total Time
- Also more failures in other groups than in KC-Viz group
  - 5 (NTK), 6 (OwlViz), 2(KC-Viz)

Other findings

- Usability
  - No much difference in usability scores
    - KC-Viz slightly better but not significant
  - However positive correlation found between experience in ontology engineering and usability score
    - That is, perceived usability probably reflects the greater ability of subjects
    - These findings suggest that the results of usability questionnaires in this kind of evaluations should be treated with much caution and ideally triangulated with other data
- Experience and Performance
  - Also significant negative correlation between ontology experience score and total time spent across the four tasks and on task 3
  - Correlation between experience of ontology engineering tools and task performance was statistically significant for task 1 and task 3 and was close to significance on overall performance
  - In sum: experts perform better regardless of the tool!

Summing up

- KC-Viz is unique in the way it uses ontology summarization to support ontology visualization and navigation
- Experimental results suggest that KC-Viz provides performance advantage to users in typical navigation tasks
- Feedback from users confirms value of KCE-based approach

Future Work

- Preliminary experiments with CropCircles show good performance with this tool in ‘pure’ topological tasks
  - This suggests the value of a very coarse-grained visualization as the basis for exploration - even though the control provided by KC-Viz appears to be more effective in supporting more complex exploration tasks
- Video Analysis also provided evidence of other features which appears to provide advantages to users
- Planned new features
  - **Coarse-grained views**, where an entire ontology is shown, even though you can't actually see the labels.
  - **Dynamic mode** - navigate in the entity browser and update graphs automatically using the current expand option
  - **Node-specific collapse/undo collapse**
  - **Ability to highlight connected links**

KNOWLEDGE MEDIA

KMi

INSTITUTE