



A Dynamic Ontology for a Dynamic Reference Work

Mathias Niepert

Cameron Buckner

Colin Allen

Indiana University, Bloomington

Outline

- Digital Tools for the Humanities
- SEP: Special Opportunities & Challenges
- Initial Ontology Design
- Algorithm Design
- Completing and Integrating the Metacontent Generation Engine
- Discussion and Future Work

Digital Tools for the Humanities

- Challenges to collaboration & integration:
 - humanities (esp. philosophy) often abstract, expansive topics, interpretively difficult & controversial
- Many ideas & “proofs of concept”, few systems ready to serve existing projects
 - opportunity afforded by the Stanford Encyclopedia of Philosophy (SEP)
 - SEP's needs: cross-referencing, listings by subject area, semantic search, faceted search, visualization, scholarly integrity (ontology vs. folksonomy)

SEP:

Special Challenges/Opportunities

- Asynchronous document submission and open-access ideals challenge traditional editorial practices
 - dynamic content requires dynamic metacontent
- Scale, authority, comprehensiveness of SEP appropriate base for automatically extracting metacontent
- SEP presents special opportunity, as “house” experts can provide feedback on automatic methods during author/editor procedures
 - but interfaces mustn't assume familiarity with goals or technical aspects of ontology design

Initial Ontology Design

- InPhO approach: hand-built ontology, semi-automatic population and extension, expert feedback/training
- Individuals: thinkers, documents, organizations, and philosophical ideas
 - e.g. Plato, *The Republic*, The Vienna Circle, logical positivism
- Idea sub-ontology most important for SEP's needs, yet most difficult to taxonomize

Two Approaches to Ontology of Ideas


- Social/structural roles vs. semantic taxonomy
 - social/structural categories e.g. concept, position, argument, distinction...
 - privileges form over content
 - not immediately useful for cross-referencing, visualization, semantic search
- Pragmatically, we started with a representation which
 - expresses semantic relevance and tracks lexical items
 - can be populated/extended using methods operating on term statistics and leveraging expert feedback

InPhO-SEP Tree (Alpha Version)

This is the page for the node [propositional attitudes](#)

This tree is the outcome of the first iteration of our cycle of expert-provided structure, statistical analysis of the articles in the [Stanford Encyclopedia of Philosophy](#), and a small amount of human feedback. Iterations of this cycle will be used to develop the tree. The approach is explained in more detail [here](#).

Click on any of the links on the left to explore the topics in the tree.

Click  icon to open SEP article on the adjacent term.

▼ philosophy

▶ [aesthetics and philosophy of art](#)

▶ [ethics](#)

▶ [feminist philosophy](#)

▶ [history of philosophy](#)

▶ [logic](#)

▶ [metaphysics](#)

▶ [philosophy of language](#)

▼ philosophy of mind

▶ [artificial intelligence](#)

▶ [consciousness and qualia](#)

▶ [consciousness and science](#)

▼ mental content

▶ [intentionality](#)

▶ [internalism and externalism](#)

▼ **propositional attitudes**

▶ [status of intentionality](#)

▶ [theories of mental content](#)


▶ [metaphysics of mind](#)


▶ [philosophy of psychology](#)


▶ [philosophy of science and the sciences](#)

▶ [social and political philosophy](#)


Instances:

 [anaphora](#)

 [structured propositions](#)

 [propositional attitude reports](#)


[eliminativism about propositional attitudes](#)

 [belief](#)

[intentional stance](#)

[desires](#)

Links:

 [@language of thought hypothesis](#)

Searches

[\[SEP\]](#) [\[noesis\]](#) [\[scholar\]](#)

[\[SEP\]](#) [\[noesis\]](#) [\[scholar\]](#)

[\[SEP\]](#) [\[noesis\]](#) [\[scholar\]](#)

[\[SEP\]](#) [\[noesis\]](#) [\[scholar\]](#)

[\[SEP\]](#) [\[noesis\]](#) [\[scholar\]](#)

[\[SEP\]](#) [\[noesis\]](#) [\[scholar\]](#)

[\[SEP\]](#) [\[noesis\]](#) [\[scholar\]](#)

Searches

[\[SEP\]](#) [\[noesis\]](#) [\[scholar\]](#)

Examples of Non-Taxonomic Relations

Thinker relations:

- **thinker properties:** birthdate, deathdate, has_profession, has_nationality
 - deathdate(Aristotle, -322), has_profession(Sartre, philosopher)
- **thinker->thinker:** teacher_of, influenced
 - teacher_of(Plato, Aristotle)

Idea relations:

- **idea properties:** idea_type (concept, position, distinction...)
 - idea_type(the intentional stance, position)

Ternary relations:

- disagreed_with, agreed_with (X disagreed with Y about Z)
 - disagreed_with(Aristotle, Plato, Theory of Forms)

Outline

- Digital Tools for the Humanities
- Dynamic Reference Works: Special Opportunities & Challenges
- Initial Ontology Design
- Algorithm Design
- Completing and Integrating the Metacontent Generation Engine
- Discussion and Future Work

Enrich & Populate Taxonomic Relations

- Maintain controlled vocabulary of terms through author feedback (TF-IDF, n-grams)
- Create “transactions” and compute co-occurrence statistics
- Frequent Pattern Mining (J-measure, semantic similarity)
- Build weighted co-occurrence graph $G = (V, E, w, l)$
- Node entropy as measure for generality

$$R(i \rightarrow j) = J(i \rightarrow j) \times \frac{H(i) - H(j)}{H_{max}} \quad H(i) = - \sum_{k \in \{j | (i, j) \in E\}} p(k|i) \log p(k|i)$$

- Provides rankings for each term node

Present rankings to authors



« myspace

- ▼ philosophy
 - ▶ aesthetics and philosophy of art
 - ▶ ethics
 - ▶ feminist philosophy
 - ▶ history of philosophy
 - ▶ logic
 - ▶ metaphysics
 - ▶ philosophy of language
 - ▼ philosophy of mind
 - ▼ **artificial intelligence**
 - ▶ computation and representation
 - ▶ computationalism
 - ▶ connectionism
 - ▶ dynamic system
 - ▶ thinking machine
 - ▶ consciousness and qualia
 - ▶ consciousness and science
 - ▶ mental content
 - ▶ metaphysics of mind
 - ▶ philosophy of psychology
 - ▶ philosophy of science and the sciences
 - ▶ social and political philosophy

SEP OntoTree

This is the page for the node **artificial intelligence**

You can navigate through the taxonomy by clicking on the topic node on the left. Clicking on a node expands it into the available subtopics. If no terms appear below, on the right side of this page, then please follow links until some do.

For each term shown below on the left, please indicate its relationship to the topic node selected (i.e., **artificial intelligence**). You may skip any items you are unsure about. For more information about what you are being asked to do, please click [here](#).

Page **1 2 3 4 5 6 7 8 9 10 11 12** [Add your own](#) | [Jump to submit button](#)

unrelated ◁ ◁ ◁ ◁ ◁ highly related

is more specific than

unrelated ◁ ◁ ◁ ◁ ◁ highly related

- is more specific than
- is more general than
- is as general as
- incomparable/either

unrelated ◁ ◁ ◁ ◁ ◁ highly related

is more specific than

Evaluation

- Finance ontology (developed within the GETESS project) as gold standard
- Small corpus of 100 documents retrieved from Wikipedia
- 1216 different concepts (i.e. terms), 1278 instances
- 1,185(92.7%) could be classified as either sub-concept-of or super-concept-of relations
- 799(67.4%) correctly and 386(32.6%) wrongly classified
- Subsumption (Sanderson&Croft): 358(28%), 255(71.2%), 103 (28.8%)

Enrich & Populate Non-Taxonomic Relations

- Internally: Hidden Markov Models and other IE methods on fragments of text (using controlled vocabulary of ideas and philosophers' names)
- Externally: Philosophers' family tree, Wikipedia (pattern matching works well for some relations), WordNet
- Problems of data integration and verification arise but author feedback can be leveraged similarly

Jean-Paul Sartre

Year of birth: Month of birth: Day of birth:

Year of death: Month of death: Day of death:

Nationality/Ethnicity:

Occupation(s):

Alternative names:

Has influenced:

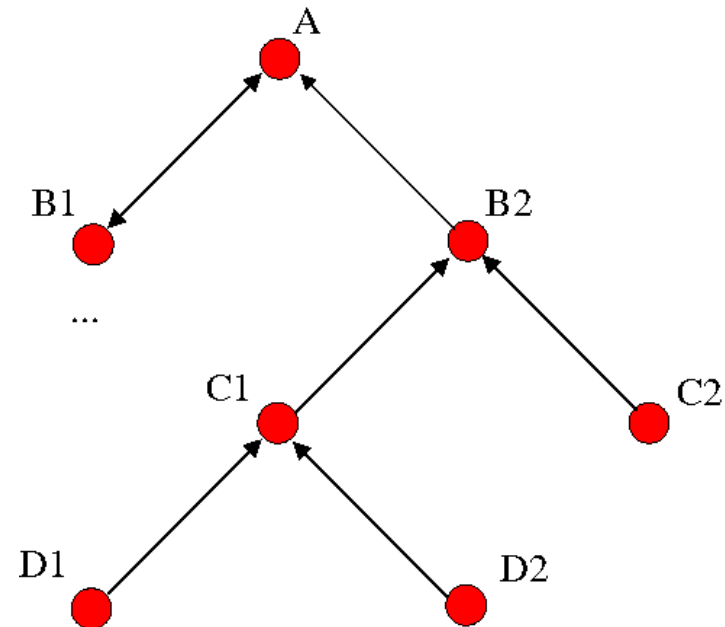
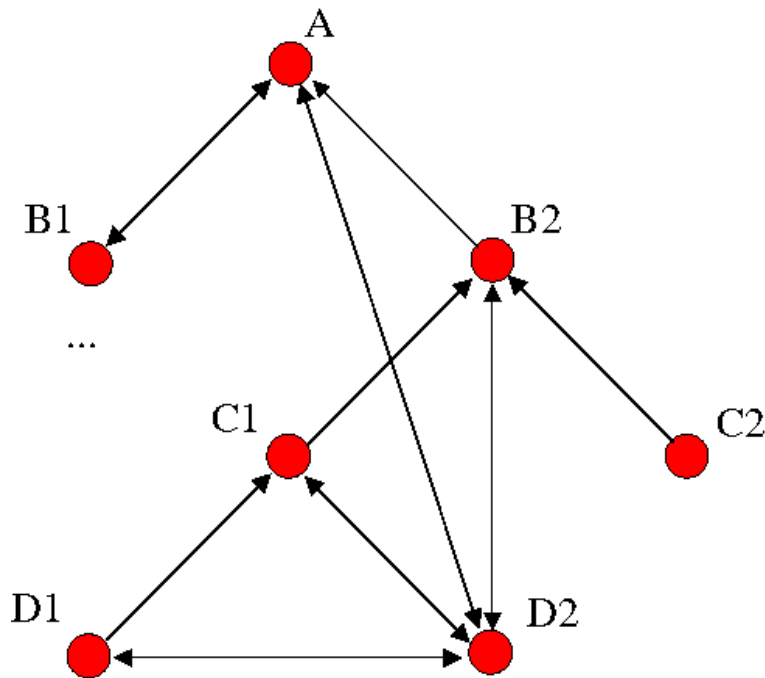
Indicate degree of influence for selected thinkers
no influence strong influence

Influenced by:

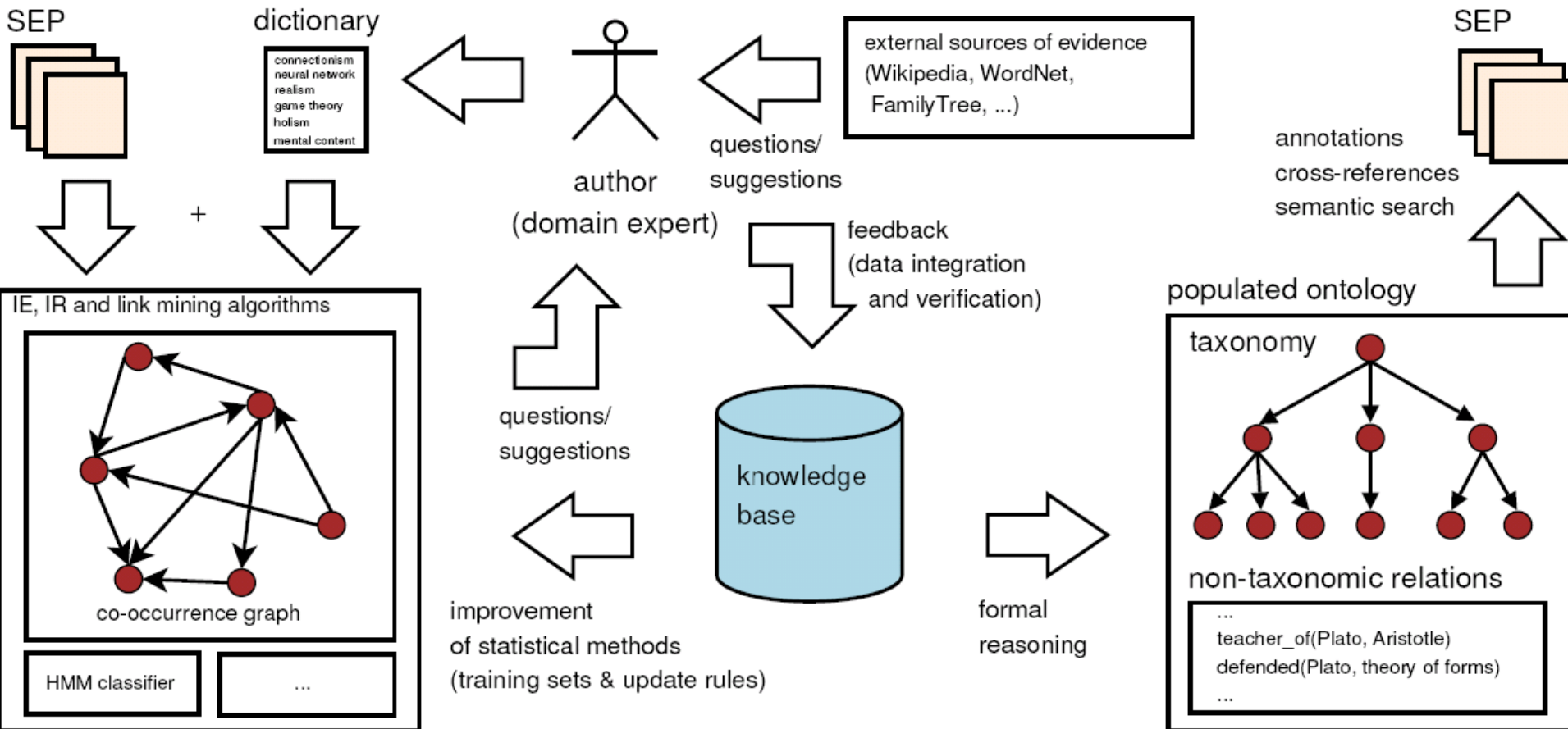
Indicate degree of influence for selected thinkers
 strong influence

Teacher of:

- Store and use the expert feedback to improve (re-train) the statistical methods
- Apply co-occurrence graph update (inference) rules
- Nonmonotonic reasoning (ASP) to infer taxonomy (with “semantic links”)



InPhO Architecture



Discussion & Future Work

- System integrates well with the editorial workflow of the SEP
- New “modules” can be added as sources for populating the “uncertain part” of the knowledge base
- Authors and editors integrate and verify the data, formal (nonmonotonic) reasoning puts the pieces together
- Visualization, Cross-Referencing, “Semantic Search”

Acknowledgments

Funded by Indiana University under the grant “New Frontiers in the Arts and Humanities” and the Cognitive Science Program

<http://inpho.cogs.indiana.edu>