

InPhOrmed Philosophy

Combining text mining and expert judgment

Colin Allen <colallen@indiana.edu>



Mathias Niepert <mniepert@indiana.edu>



Cameron Buckner <cbuckner@indiana.edu>



1



Goals

- To build and maintain a “dynamic ontology” for the discipline of philosophy
- To deploy the InPhO in a variety of Digital Philosophy applications.



Digital Tools for the Humanities



2

Cross references for Functionalism

| 1 | 2 | 3 | 4 | 5 | Article | 1 | Biographical Article |
|-------------------------------------|----|----|----|----|--|--------------------------|-----------------------|
| <input type="checkbox"/> | 8 | 2 | 1 | 10 | zombies | <input type="checkbox"/> | 1 Ludwig Wittgenstein |
| <input checked="" type="checkbox"/> | 3 | | | 1 | behaviorism | <input type="checkbox"/> | 2 Gilbert Ryle |
| <input checked="" type="checkbox"/> | 3 | 2 | 11 | | identity theory of mind | <input type="checkbox"/> | 3 Donald Davidson |
| <input checked="" type="checkbox"/> | 12 | | 8 | 2 | physicalism | <input type="checkbox"/> | 4 Alan Turing |
| <input type="checkbox"/> | 1 | | 2 | | pain | <input type="checkbox"/> | 5 Franz Brentano |
| <input checked="" type="checkbox"/> | 2 | | 3 | | qualia | <input type="checkbox"/> | 6 John Locke |
| <input type="checkbox"/> | | 1 | 4 | | inverted qualia | | |
| <input checked="" type="checkbox"/> | 4 | | 5 | | belief | | |
| <input type="checkbox"/> | 6 | | 6 | | identity | | |
| <input checked="" type="checkbox"/> | 5 | | 9 | | multiple realizability | | |
| <input checked="" type="checkbox"/> | 15 | | 1 | | dualism | | |
| <input type="checkbox"/> | 5 | | 12 | | concepts | | |
| <input type="checkbox"/> | 11 | | 7 | | dispositions | | |
| <input checked="" type="checkbox"/> | 18 | 12 | | | mental causation | | |
| <input type="checkbox"/> | 4 | | | | anomalous monism | | |
| <input type="checkbox"/> | 6 | | | | chinese room argument | | |
| <input type="checkbox"/> | 7 | | | | property | | |
| <input type="checkbox"/> | 7 | | | | game theory and ethics | | |
| <input type="checkbox"/> | 8 | | | | representational theories of consciousness | | |
| <input type="checkbox"/> | 9 | | | | color | | |
| <input type="checkbox"/> | 9 | | | | teleological theories of mental content | | |
| <input type="checkbox"/> | 10 | | | | consciousness | | |
| <input type="checkbox"/> | 10 | | | | externalism about mental content | | |
| <input checked="" type="checkbox"/> | 11 | | | | knowledge argument | | |
| <input type="checkbox"/> | 13 | | | | reference | | |
| <input type="checkbox"/> | 14 | | | | descriptions | | |
| <input type="checkbox"/> | 16 | | | | memory | | |
| <input type="checkbox"/> | 17 | | | | propositions | | |
| <input type="checkbox"/> | 19 | | | | evidence | | |
| <input type="checkbox"/> | 20 | | | | species | | |
| <input checked="" type="checkbox"/> | | | | | computational theory of mind | | |
| <input checked="" type="checkbox"/> | | | | | mental representation | | |
| <input checked="" type="checkbox"/> | | | | | turing machine | | |

Idea Sources:

- 1 - SEP article occurrence
- 2 - Hyponymy
- 3 - User evaluations
- 4 - Jweight
- 5 - Taxonomic relations

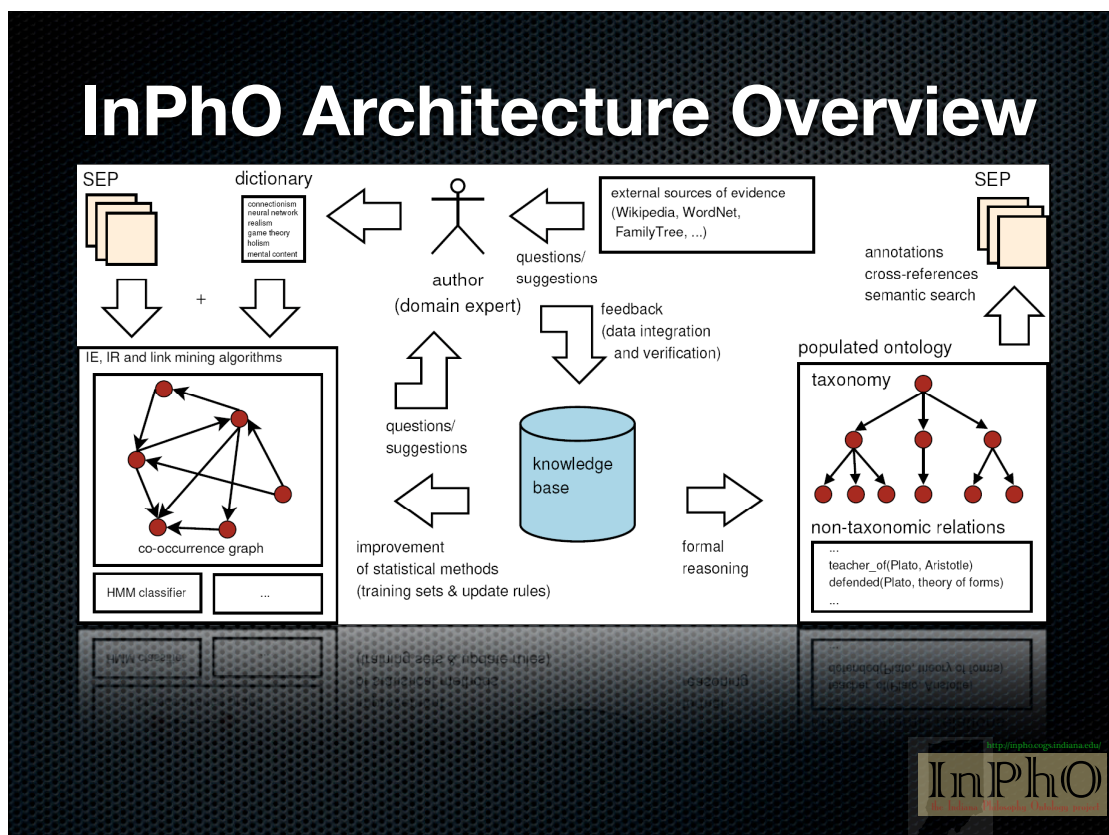
Thinker Sources:

- 1 - Idea-thinker jweight

☒ computational theory of mind
☒ mental representation
☒ turing machine

Submit References

3



4

SEP Data

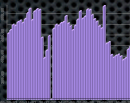
1289 authors (+12)

114 subject editors (+1)

1037 published
entries (+20)

12.05 million words
(+250,000)

600-850K articles
accessed/week



Stanford Encyclopedia of Philosophy

STANFORD ENCYCLOPEDIA OF PHILOSOPHY

Principal Editor: Edward N. Zalta

Table of Contents

U.S.A. Main Site: Stanford University
Center for the Study of Language and Information

Australia Mirror: University of Sydney
Sydney Electronic Text and Image Service

Netherlands Mirror: University of Amsterdam
Institute for Logic, Language and Computation

United Kingdom Mirror: University of Leeds
Subject Center for Philosophical & Religious Studies

Current Operations Are Supported By:

- The Offices of the Provost, the Dean of Research, and the Dean of Humanities and Science, Stanford University
- The SEP Library Fund: with contributions from the National Endowment for the Humanities and the membership dues of academic and research libraries that have joined SEPIA.
- The Friedrich Hayek Fund & The SEP Fund: with contributions from private donors and individual readers.

The SEP gratefully acknowledges founding support from the National Endowment for the Humanities, the National Science Foundation, The American Philosophical Association/Pacific Division, The Canadian Philosophical Association, and the Philosophy Documentation Center. Current fundraising efforts are being supported with a grant from The William and Flora Hewlett Foundation.

LIBRARY OF CONGRESS CATALOG DATA: ISSN 1095-5054
COPYRIGHT © 2008 BY THE METAPHYSICS RESEARCH LAB
CENTER FOR THE STUDY OF LANGUAGE AND INFORMATION, STANFORD UNIVERSITY, STANFORD, CA 94305

5

Other data sources

THE UNIVERSITY OF TEXAS AT AUSTIN

Department of Philosophy, University of Texas at Austin, 1 University Station C3500, Austin, TX 78712
Josh Dever; Woggonen 4108; 512-471-5511; FAX: 512-471-4605; dever@mail.utexas.edu

The Philosophy Family Tree

Tree

Statistics

Currently indexing 8946 Philosophers

Nguyen-D

WordNet a lexical database for the English language

Lists of philosophy topics

From Wikipedia, the free encyclopedia

Philosophy topics

Alphabetical index: (A-C) (D-H) (I-L) (M-N) (O-P) (Q-R) (S-T) (U-V) (W-X) (Y-Z)

Philosophers: (A-C) (D-H) (I-L) (M-N) (O-P) (Q-R) (S-T) (U-V) (W-X) (Y-Z)

Lists | Basic topics | Glossary of philosophy

For philosophy topics arranged alphabetically, see the philosophy lists include:

Contents [hide]

- General
- Areas of philosophy
- Timelines of philosophical progression
- See also

General

- Glossary of philosophical isms
- List of basic philosophy topics
- List of philosophies
- Lists of philosophers
- List of philosophical questions
- List of philosophical topics (index)

Areas of philosophy

- List of ethics topics
- List of topics in metaphysics
- List of topics in epistemology
- List of topics in logic
- List of topics in aesthetics
- List of topics in philosophy of life
- List of topics in philosophy of mind
- List of topics in Ancient philosophy
- List of philosophy journals
- List of belief systems

6

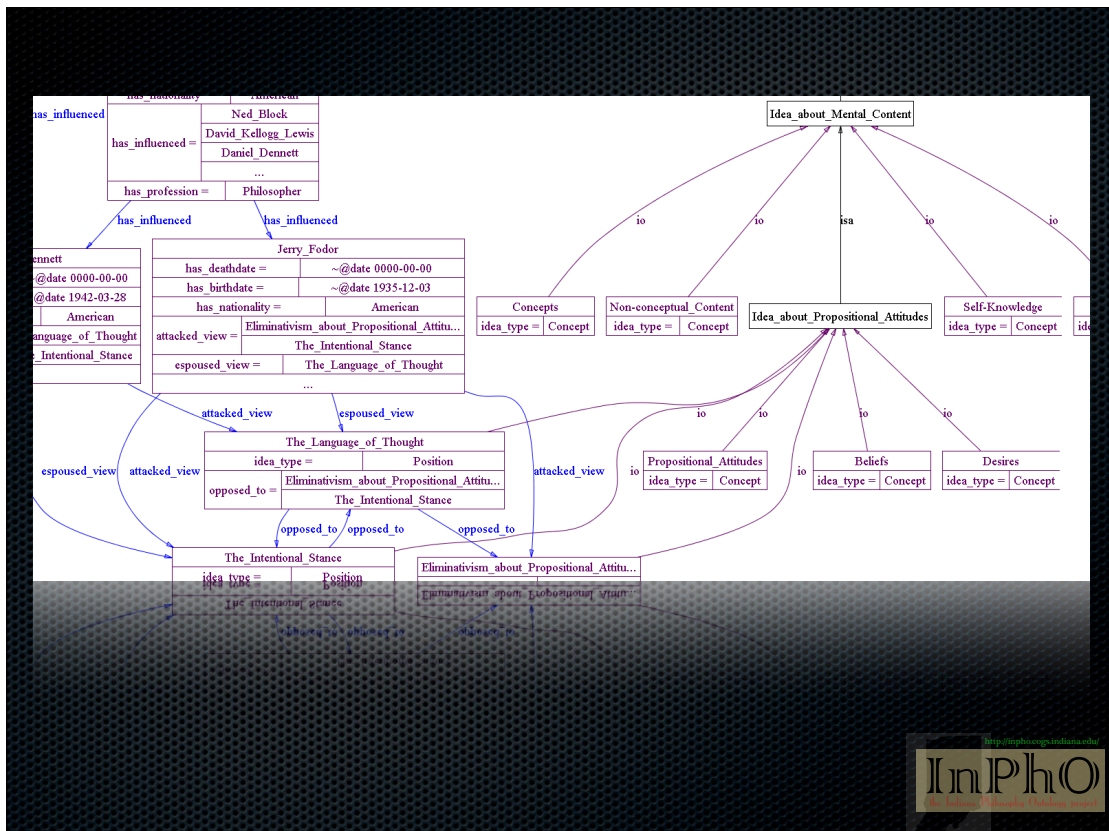
Wanted: Structured Data Out

Ontology: formal, machine-readable specification of the types of entities in a domain and relationships between them

| Table 1: Initial Property List | |
|---------------------------------------|--|
| <i>Thinker Properties</i> | Died.on, Born.on , Spoke.language , Nationality |
| <i>Thinker-Thinker Relations</i> | Teacher.of, Influenced, Criticized, Defended, Dissertation_Advisor.of, Discoursed.with |
| <i>Idea Properties</i> | Idea.type [concept, position, etc.] |
| <i>Thinker-Document Relations</i> | Wrote, Edited |
| <i>Thinker-Organization Relations</i> | Member.of, Studied.at |
| <i>Thinker-Idea Relations</i> | Worked.on (problem), Created.view, Attacked.view, Espoused.view, Aware.of |
| <i>Idea-Idea Relations</i> | Opposed.to, Commits.to (idea1 commits one to idea2) |
| <i>Document-Document Relations</i> | Published.in (article in journal/book) |
| <i>Document-Idea Relations</i> | Discusses |
| <i>Ternary Relations</i> | Disagreed.with (ThinkerX disagreed with ThinkerY on IdeaZ) |



7



8

Bridging the Data-Metadata Gap

- Two “extremes”:
 - Hire experts to design & maintain an ontology
 - Problems: labor-intensive, expensive, depends on “double” experts
 - Tagging approaches, folksonomies
 - Problems: may not meet academic standards; noisy

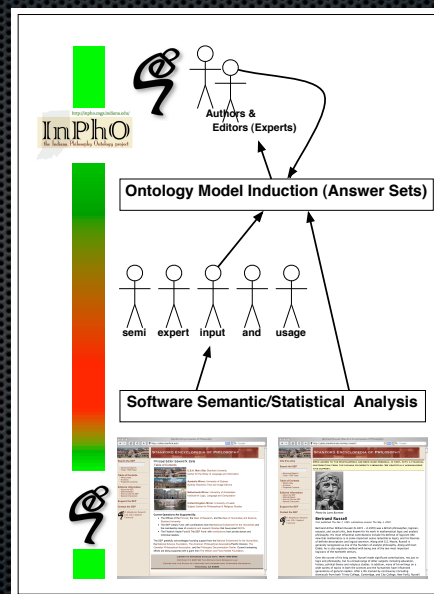


9

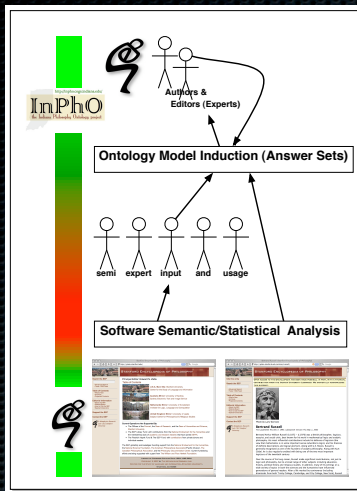
Building ontologies: A third way

- Stratified collaboration
 - Expert feedback
 - software
 - General feedback
 - software
 - Expert-written content

The InPhO “layer cake”



10

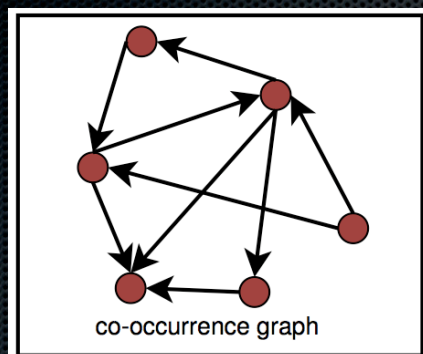


- Experts are **busy** people. Experts don't want to be bothered with **garbage**. Experts don't like their **hard work messed up** by amateurs.
- Knowledgeable amateurs often have more **time** and **motivation** to fix things, but they are **rare**. They don't like having their **hard work messed up** by the clueless either.
- Well-intentioned amateurs are **plentiful** and **motivated** to donate their **time**, But they make **mistakes**.
- Software has lots of **time**, has **no motivation problems** but is **clueless**.

<http://inpho.org/midiana.edu/>
InPhO

11

Two networks



12

Software Layer 1

- Is it possible to detect hyper/hyponymy statistically?
 - If t_1 is a hypernym of t_2 then
 - t_1 is semantically similar to t_2
 - t_1 is more general than t_2 with respect to a taxonomy
 - Probabilistic J-measure widely used to estimate the semantic similarity between two terms:

$$J(i \rightarrow j) =$$

$$p(i) \left(p(j|i) \log \frac{p(j|i)}{p(j)} + (1 - p(j|i)) \log \frac{1 - p(j|i)}{1 - p(j)} \right)$$



13

Semantic similarity

- We build a directed and weighted co-occurrence graph $G = (V, E)$ in which each node represents a term in our set of keywords. An edge between two terms t_1 and t_2 indicates that the terms co-occur in the encyclopedia at least once and the weight of the edge is a measure of their semantic similarity
- By iterating over all documents in the encyclopedia and counting their term (co-)occurrences we can estimate the probabilities $p(t_i)$, $p(t_j)$, $p(t_i, t_j)$, and thus $p(t_i|t_j)$ for all terms t_i, t_j , with respect to a unit of text.
- Currently we consider a document and a one sentence sliding window as units of text; i.e. two co-occurrence records are created, one for the sentence level and one for the document level. The latter is projected into the former by treating an entire document as one large sentence, giving us more co-occurrences at the cost of possibly including some connected but unrelated terms in our graph.



14

Hyper/hyponymy

- We hypothesize encyclopedias are “balanced” – terms representing more general categories tend to co-occur with more terms in the encyclopedia’s text.
- Normalized node in degree will usually be a good measure for the generality of category, but we anticipate that entropy is an even better approximation of generality because it not only takes into account the in-degree of the node but also how evenly its adjacent nodes are conditionally distributed.
- Node entropy provides a measure for generality that can be used to rank hypernym/hyponym candidates via the “R-measure” (Niepert et al. 2007):

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



15

| ← T → | ID | ante_name | ante_id | cons_name | cons_id | confidence | jweight | weight |
|--------------------------|---------|--------------------------------|---------|------------------------------------|---------|------------|------------|-----------------------|
| <input type="checkbox"/> | 4118455 | interpretations of probability | 1212 | probability or statistics | 984 | 0.462572 | 0.00314345 | 0.00055116226184344 |
| <input type="checkbox"/> | 4194454 | mind | 2194 | body | 2163 | 0.0964258 | 0.0031342 | 6.21925329774964e-05 |
| <input type="checkbox"/> | 4155520 | identity | 1581 | personal identity and ethics | 1666 | 0.0791403 | 0.00312541 | -0.000765116937470417 |
| <input type="checkbox"/> | 4192979 | expression and idiom | 875 | reference and denotation | 876 | 0.162823 | 0.00312527 | -2.18815713364815e-05 |
| <input type="checkbox"/> | 4045690 | memory | 1850 | epistemological problems of memory | 1299 | 0.137268 | 0.00312259 | -0.000948000268064484 |
| <input type="checkbox"/> | 3975746 | existence | 1545 | existence of god | 2013 | 0.0504892 | 0.00311267 | -0.000175873536732548 |
| <input type="checkbox"/> | 4192590 | belief | 921 | iustification | 2195 | 0.0840042 | 0.00310113 | -4.43858128130508e-06 |

Table 2: Top 10 ranked hyponym candidates for the term-node *science* by *R*-value.

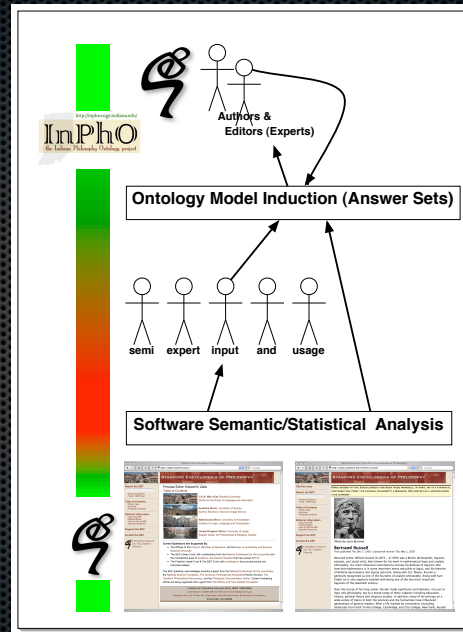
| rank | node j | in-deg. | node i | in-deg. | J | H(i)-H(j) | R | P(i j) |
|------|----------------------|---------|---------|---------|--------|-----------|--------|--------|
| 1 | cognitive science | 409 | science | 823 | 0.5046 | 0.0502 | 0.0253 | 1.0 |
| 2 | computer science | 294 | science | 823 | 0.2101 | 0.0653 | 0.0137 | 1.0 |
| 3 | social science | 412 | science | 823 | 0.4225 | 0.0283 | 0.0119 | 1.0 |
| 4 | conservation biology | 61 | science | 823 | 0.0108 | 0.3308 | 0.0036 | 0.2097 |
| 5 | physics | 688 | science | 823 | 0.1872 | 0.0171 | 0.0032 | 0.2073 |
| 6 | laws of nature | 460 | science | 823 | 0.0424 | 0.0642 | 0.0027 | 0.1796 |
| 7 | mathematics | 633 | science | 823 | 0.1575 | 0.016 | 0.0025 | 0.2143 |
| 8 | biology | 522 | science | 823 | 0.1127 | 0.0217 | 0.0024 | 0.2323 |
| 9 | molecular biology | 193 | science | 823 | 0.016 | 0.1496 | 0.0024 | 0.1789 |
| 10 | neuroscience | 308 | science | 823 | 0.0473 | 0.0491 | 0.0023 | 0.3419 |



16

Leveraging Expertise

- Simple question interface to gather feedback on statistically generated “hypotheses”
- Automated (nonmonotonic) reasoning to put the pieces together (Answer Set Programming)



17

This is the page for the node **philosophy of cognitive science**

You can navigate through the Idea tree by clicking on the topic node links 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., **philosophy of cognitive science**). You may skip any items you are unsure about. For more information about what you are being asked to do, please click [here](#).

- How should I decide the relatedness or relative generality of two ideas? (see 1-4)
- How hard should I think about idea pairs which seem odd?
- Can I review/revise my feedback?

Page **1** **2** **3** **4** **5** **6** **7** **8** **9** **10** Add your own | Jump to submit button

unrelated ○ ○ ○ ○ ○ highly related [reset] [SEP] [Google SEP]

mental imagery incomparable/either philosophy of cognitive science

☐ This is not a philosophical idea. ?

unrelated ○ ○ ○ ○ ○ highly related [reset] [SEP] [Google SEP]

folk psychology is more specific than is more general than is as general as incomparable/either philosophy of cognitive science

☐ This is not a philosophical idea. ?

unrelated ○ ○ ○ ○ ○ highly related [reset] [SEP] [Google SEP]

connectionism is more specific than philosophy of cognitive science

☐ This is not a philosophical idea. ?

18

Thinker

MyInPhO (colallen)

Search database for: Alan Turing

Alan Turing

Year of birth: 1912 Month of birth: June Day of birth: 23
Year of death: 1954 Month of death: June Day of death: 7

Nationality/Ethnicity:
Occupation(s): Cryptographer, Logician, Mathematician
Alternative names:

Has influenced: Ned Block, Noam Chomsky, Daniel Dennett, Hilary Putnam
Indicate degree of influence for selected thinkers
no influence ☐ ☐ ☐ ☐ strong influence

Influenced by: Bertrand Russell
Indicate degree of influence for selected thinkers
no influence ☐ ☐ ☐ ☐ strong influence

Teacher of:
Indicate if selected thinkers were Alan Turing's students
no ☐ ☐ yes

Student of: Alonzo Church
Indicate if selected thinkers taught Alan Turing
no ☐ ☐ yes

Ternary statements:
Indicate strength of selected statement
weak ☐ ☐ ☐ ☐ strong

Alan Turing on

InPhO
the Indiana Philosophy Ontology project

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

19

Software 2

- Is it possible to use possibly conflicting evaluations?
- Yes: Non monotonic reasoning.
- We use answer-set programming.

```

mg(X, Y) ← ms(Y, X) (1)
ms(X, Y) ← mg(Y, X) (2)
ic(X, Y) ← ic(Y, X) (3)
si(X, Y) ← si(Y, X), i ∈ {0, ..., 4} (4)
mg(X, Y) ← mg(X, Z), mg(Z, Y) (5)
ms(X, Y) ← ms(X, Z), ms(Z, Y) (6)
ic(X, Y) ← ms(X, Y), mg(X, Y), si(X, Y), i ∈ {3, 4} (7)
desc(X, Y) ← isa(X, Y), class(X), class(Y) (8)
desc(X, Z) ← desc(X, Y), desc(Y, Z), class(X), (9)
class(Y), class(Z)
pins(X, Y) ← si(X, Y), ms(X, Y), not class(X), (10)
class(Y), not ic(X, Y), i ∈ {3, 4}
plink(X, Y) ← si(Y, X), ms(Y, X), not desc(Y, X), (11)
not ic(Y, X), class(X), class(Y), i ∈ {3, 4}
plink(X, Y) ← si(X, Y), ic(X, Y), not desc(Y, X), class(X) (12)
nins(X, Y) ← pins(X, Z), desc(Z, Y), not class(X), (13)
class(Y), class(Z)
nins(X, Y) ← pins(X, Z), plink(Y, Z), not class(X), (14)
class(Y), class(Z)
instance-of(X, Y) ← pins(X, Y), not nins(X, Y) (15)
nlink(X, Y) ← instance-of(Y, Z), plink(X, Z), class(X), (16)
not class(Y), class(Z)
nlink(X, Y) ← instance-of(Y, Z), desc(Z, X), class(X), (17)
not class(Y), class(Z)
nlink(X, Y) ← instance-of(Y, X), class(X), not class(Y) (18)
nlink(X, Y) ← plink(X, Z), desc(Y, Z), class(X), (19)
class(Y), class(Z)
nlink(X, Y) ← isa(X, Z), isa(Y, Z), class(X), (20)
class(Y), class(Z)
links-to(X, Y) ← plink(X, Y), not nlink(X, Y)

```

InPhO
http://inpho.cogs.indiana.edu/

20

Answer Set Programming

- Three parts:
 - Signature: predicate symbols (e.g., desc) and set of terms (here: terms referring to ideas in Philosophy)
 - Declaration: Set of feedback facts, (e.g., more-specific(Neural Network, Connectionism)) and the facts given by the existing ontological structure (e.g., is-a(Thinking Machines, Artificial Intelligence))
 - Regular Part (set of rules)



21

Answer Set Programming Conflicting Feedback

- Conflicting feedback is possible!
- Modeled using predicate ic (inconsistent):
 - $ic(X, Y) :- ms(X, Y), mg(X, Y).$
- Can be used to model “semantic links” between incomparable ideas:
 - $plink(X, Y) :- s4(X, Y), ic(X, Y), not\ desc(X, Y), class(X).$



22

Quality of response

- Is it possible to use feedback according to quality?
 - If X is an author or editor at the SEP, then X is an expert in subject areas a, b, \dots
 - If Y provides feedback in area a that is well correlated with experts in a , then Y 's feedback about edge E in a may be trusted in the absence of contrary expert feedback.

Your recent InPhO contributions:

Philosophers:
Laozi
Desiderius Erasmus
Nishitani Keiji

Ideas (115 evaluations):
social and political philosophy
japanese philosophy
chinese philosophy

Statistics: [?]
Rel.agree: 85%
Rel.StdDev: 0.28
Gen.agree: 80%



23

Managing Expertise

- Is it possible to use feedback according to quality?
 - If X is an author or editor at the SEP, then X is an expert in subject areas a, b, \dots
 - If Y provides feedback in area a that is well correlated with experts in a , then Y 's feedback about edge E in a may be trusted in the absence of contrary expert feedback.

```
%stratified input predicates are named same as unstratified counterparts but
with "i" suffix
rating(0).
rating(1).
rating(2).
rating(3).
rating(4).

%process expert-level-stratified input versions of predicates into those used by
program
%inverse
msi(X, Y, A) ← mgi(Y, X, A), rating(A).
mgi(X, Y, A) ← msi(Y, X, A), rating(A).

%similarity symmetry
p4i(X, Y, A) ← p4i(X, X, A), rating(A).
p3i(X, Y, A) ← p3i(Y, X, A), rating(A).

%incomparable symmetry
ici(X, Y, A) ← ici(Y, X, A), rating(A).

%evidence against similarity ratings if contradicted by those at a higher level
np4(X, Y, A) ← p4i(X, Y, A), p0i(X, Y, B), B>A, rating(A), rating(B).
np4(X, Y, A) ← p4i(X, Y, A), p1i(X, Y, B), B>A, rating(A), rating(B).
np3(X, Y, A) ← p3i(X, Y, A), p0i(X, Y, B), B>A, rating(A), rating(B).
np3(X, Y, A) ← p3i(X, Y, A), p1i(X, Y, B), B>A, rating(A), rating(B).

%if no evidence against the similarity at that level, allow to pass through the
"filter"
p4(X, Y) ← p4i(X, Y, A), not np4(X, Y, A), rating(A).
p3(X, Y) ← p3i(X, Y, A), not np3(X, Y, A), rating(A).

%allow lower-level generality evaluations to pass through if not contradicted by
a higher level
%mg(X, Y) ← mgi(X, Y, A), rating(A), rating(B), not msi(X, Y, B), B>A.
%ms(X, Y) ← msi(X, Y, A), rating(A), rating(B), not msi(X, Y, B), B>A.

%evidence against the generality ratings at one level if contradicted by those at
a higher level
nmg(X, Y, A) ← mgi(X, Y, A), msi(X, Y, B), B>A, rating(A), rating(B).
nms(X, Y, A) ← msi(X, Y, A), mgi(X, Y, B), B>A, rating(A), rating(B).

%if no evidence against the generality at that level, allow to pass through the
"filter"
mg(X, Y) ← nmg(X, Y, A), not nms(X, Y, A), rating(A).
ms(X, Y) ← nms(X, Y, A), not nmg(X, Y, A), rating(A).
```



24

Representing Philosophy

Three Models

- Wiki — Power to the people! The world is flat!
- Peer reviewed — Experts know best!
Mountaintop sanctuaries (SEP, “Formal” Ontology)
- Stratified — From each according to ability! A complex landscape (InPhO)



25

Site under development! Please send comments or feedback to inpho@indiana.edu

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

Project Director
Colin Allen

Graduate Assistants | Programmers | Brains
Cameron Buckner • Mathias Niepert

Server Administration | Technical Support & Programming
Dr. Ruth Eberle • Jaimie Murdock

Data and Web Interfaces
Tarun Gangwani

Funding
National Endowment for the Humanities Digital Humanities Initiative
IU New Frontiers in the Arts and Humanities Program

Affiliated Projects and Researchers
Prof. Tony Beavers, University of Evansville • Michele Pasin, Open University, UK
Stanford Encyclopedia of Philosophy • Noesis: Philosophical Research Online

About the project
« myinpho (login)

Tools (experimental)
Browse the taxonomy

Articles & Papers
2007
• JCDL (full paper)
• APA Newsletter (brief note)
2008
• FLAIRS (full paper)
• Draft for Synthese special issue

Download the Taxonomy (alpha version!)
OWL format

Any views, findings, conclusions or recommendations expressed in this website do not necessarily represent those of the National Endowment for the Humanities.

Cognitive Science Program, 819 Eigenmann, 1910 E. 10th St.,
Indiana University, Bloomington, IN 47406-7512 USA

| Comments? |

Copyright 2007, The Trustees of Indiana University
Copyright Complaints

INDIANA UNIVERSITY
BLOOMINGTON

26