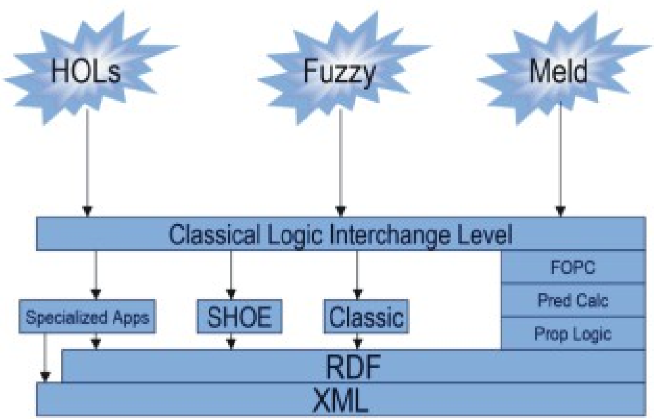
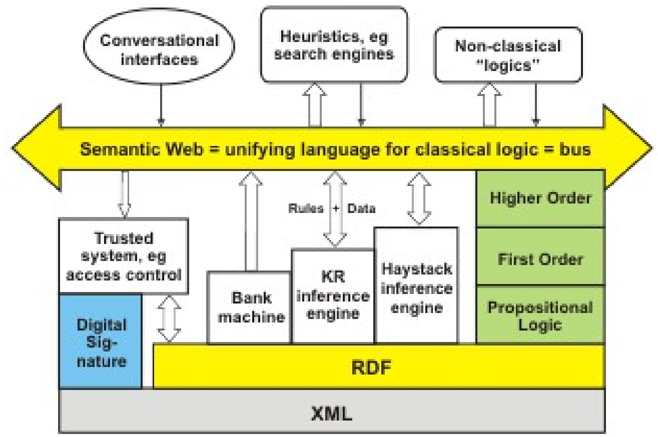
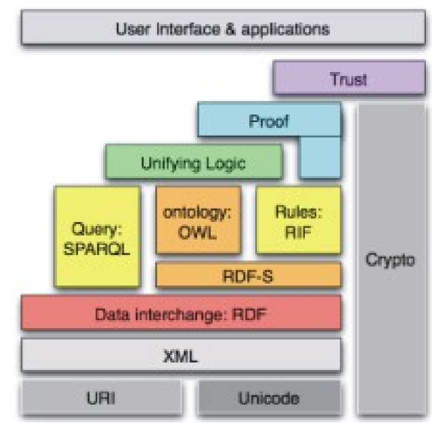
Language, Ontology, And the Semantic Web







John F. Sowa

1 September 2020

Language, Ontology, and the SW

Abstract: ***In 2000, Tim Berners-Lee proposed a vision for the Semantic Web that was more ambitious than the results delivered in 2005. Research in the past 15 years produced advanced technology in artificial intelligence, language processing, and reasoning methods, both formal and informal. But many systems are proprietary, incompatible with one another, and too complex for widespread adoption. Among the most important requirements, trusted systems were never adequately implemented. This talk surveys promising developments and suggests ways of adapting them to the Semantic Web.***

Contents:

1. Semantic Web From 2000 to 2005

In 2000, Tim Berners-Lee wrote an ambitious proposal.*

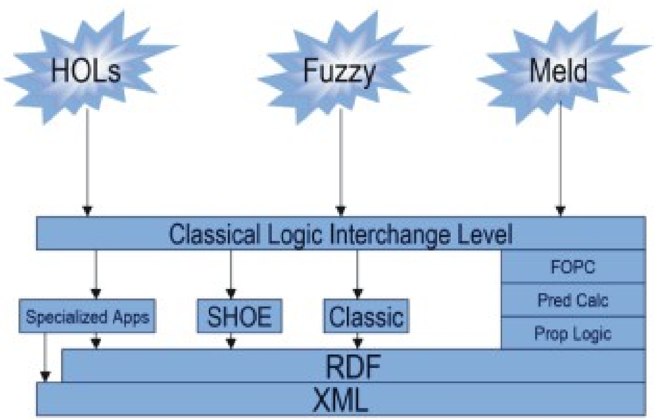
- The Semantic Web "as an interchange bus for on-line data."
- RDF as a simple language for exchanging "raw data" among "heterogeneous systems."
- SWeLL (Semantic Web Logic Language) "extends RDF by including negation and explicit quantification."
- SWeLL should represent first-order and higher-order logic and pair "simple, predictable, reliable systems with complex, unpredictable, heuristic systems."
- But the tools delivered in 2005 were more limited.

Goal: Implement Tim's vision with a new generation of tools.

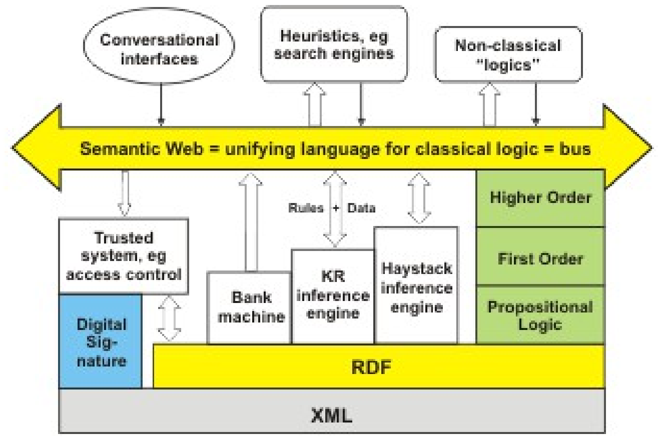
- More advanced methods for language, learning, and reasoning.
- Better methods for building trust and ensuring security.

* For the original documents, see http://jfsowa.com/ikl/

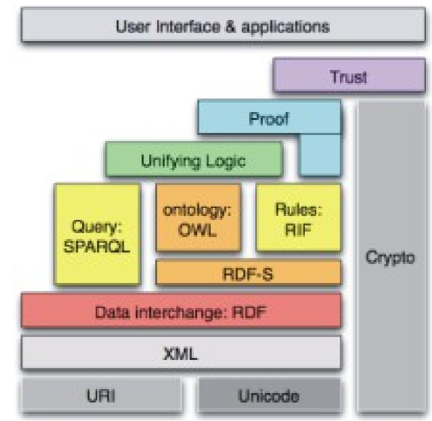
Semantic Web “Layer Cakes”



DAML Requirements (2000)



Winning Proposal (2000)



Final Report (2005)

The proposal was more ambitious than what was delivered.

- New AI technology has been developed in the past 15 years.
- Advanced applications have gone far beyond the tools of 2005.
- The new technology should be more widely available.
- It should be easy to learn, easy to use, and upward compatible.

In the diagram, the large yellow arrow is the SWeLL bus:

- Semantic Web = unifying language for classical logic = bus.

Two Examples of Advanced AI

Google's knowledge graphs (KGs) are represented in RDF.

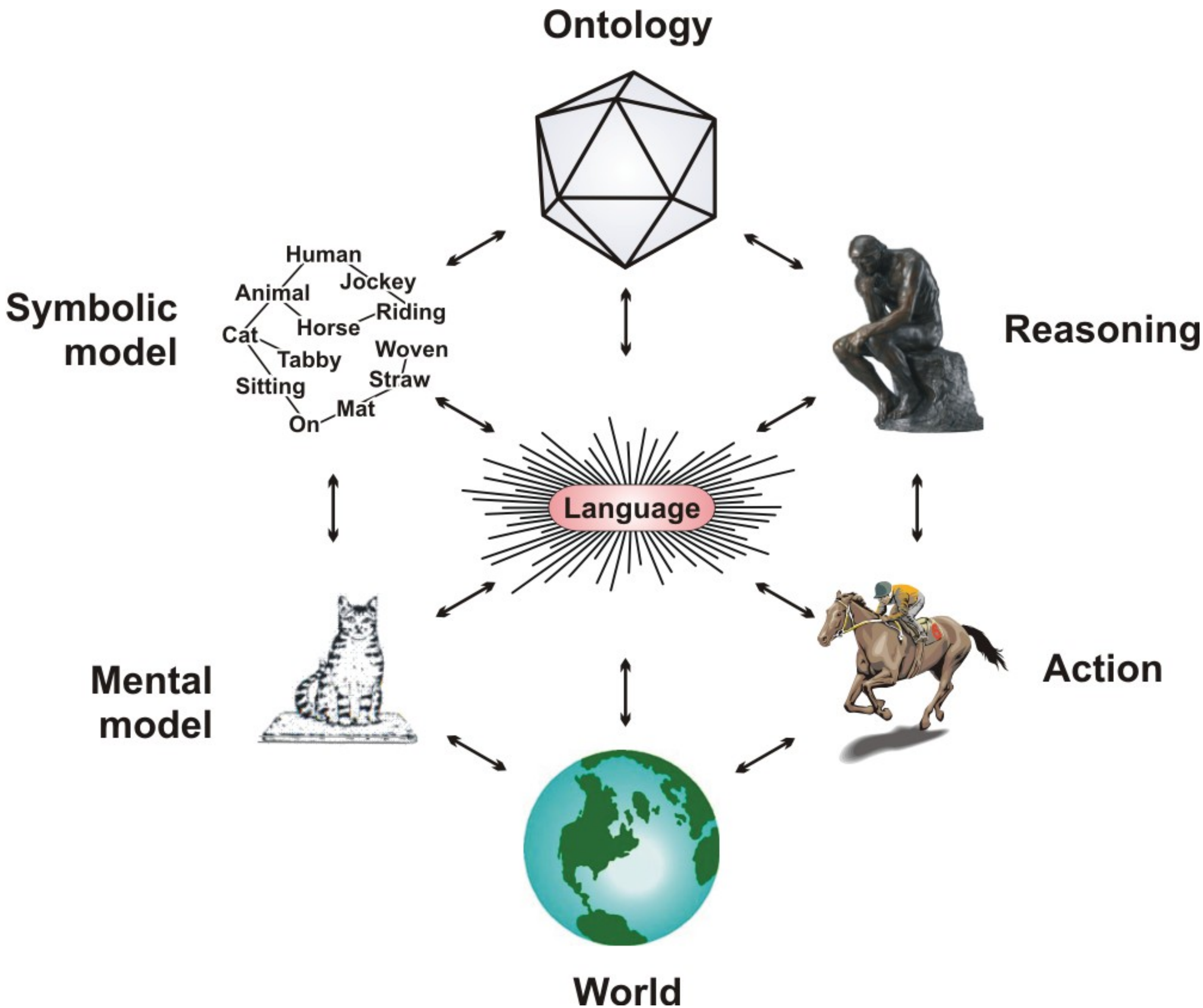
- **DBpedia and other freely available resources provide the data.**
- **Google added AI methods for learning and reasoning with KGs and deriving new KGs from documents.**

For the Jeopardy challenge, IBM Watson also used DBpedia.*

Watson added a wide range of AI technology: English parsers, question classification, question decomposition, automatic source acquisition and evaluation, entity and relation detection, logical form generation, statistics, machine learning, knowledge representation, and several methods of reasoning.

Goals for the future: Automated and semi-automated tools to make such systems easier to design, cheaper to build, and more reliable.

*** See https://www.aaai.org/Magazine/Watson/watson.php**



RDF can represent symbolic models that are directly related to ontology, language, and mental models. Tools for logic and language can relate them to reasoning, action, and the world.

Implementing the Hexagon

The corners of the hexagon represent aspects of knowledge.

1. The world is everything we encounter in space and time.
2. Mental models represent everything we experience or imagine.
3. Symbolic models consist of words related by words to other words.
4. Ontology is a catalog of words and the kinds of things they refer to.
5. Reasoning includes all our ways of thinking about anything.
6. Action is what our thinking leads us to do in and on the world.

Natural languages represent conscious knowledge.

- They can represent and relate all six corners of the hexagon.
- Every artificial language, notation, or diagram is a simplified or stylized version of something that could be said in a natural language.
- But the nervous system contains an enormous amount of unconscious knowledge that supports the basic operations of the human body.

Challenge: Implement AI tools to support all the above.

2. Interoperability

DOL is a standard for integration and interoperation among distributed ontologies, models, and specifications (OMS). *

- **UML and the Semantic Web logics are supported by DOL.**
- **DOL tools can relate anything specified by those logics or an open-ended variety of others.**
- **That includes the notations for representing legacy software and the latest technologies of the 21st century.**

DOL is formally defined by logic and mathematics.

- **Logic is essential for guaranteeing precision.**
- **DOL can integrate heterogeneous OMS by relating the logics that specify them. Common Logic (CL) is one of the most general.**
- **But people may continue to use any notations they prefer.**

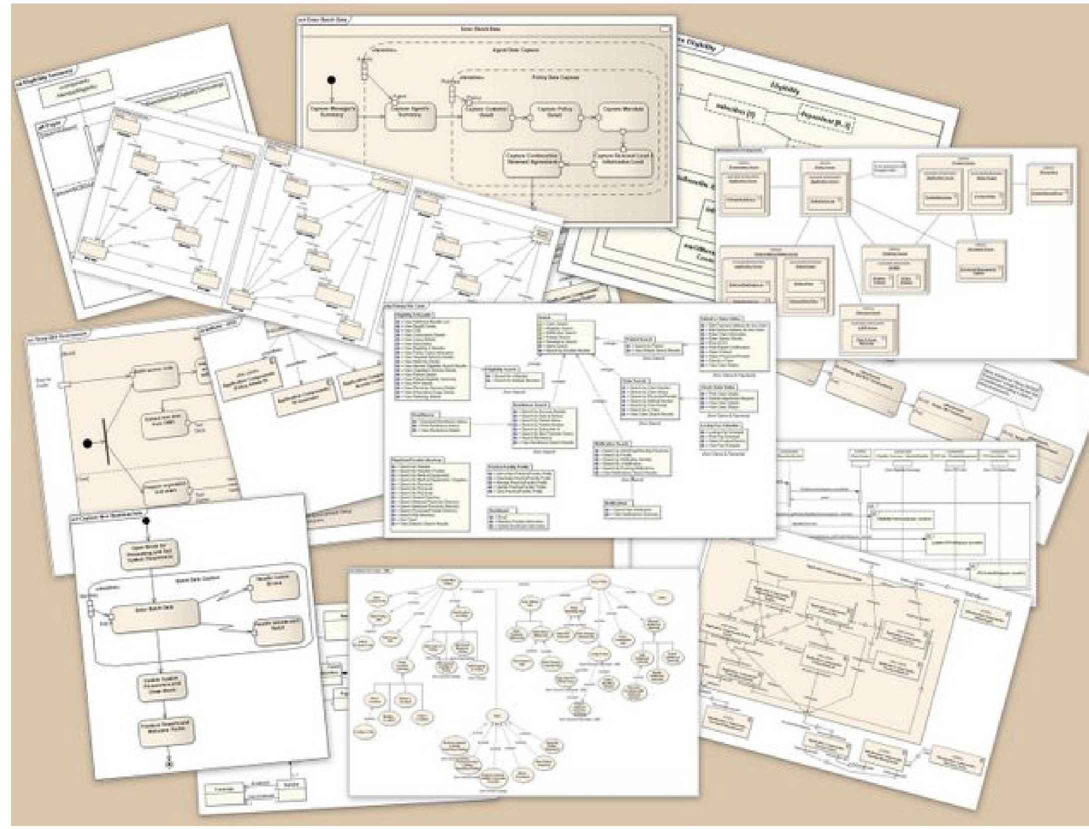
*** OMG Standard for DOL: Distributed Ontology, Modeling, and Specification Language: https://www.omg.org/spec/DOL/1.0**

Unified Modeling Language (UML)

A family of diagrams for representing database and computer system designs.

Originally specified as informal notations without a precise definition in logic.

The Object Management Group (OMG) standardized formal UML by definitions stated in Common Logic.*



By mapping UML diagrams and SW logics to CL, DOL can facilitate data sharing among applications in any field.

* See https://www.omg.org/spec/FUML/1.4

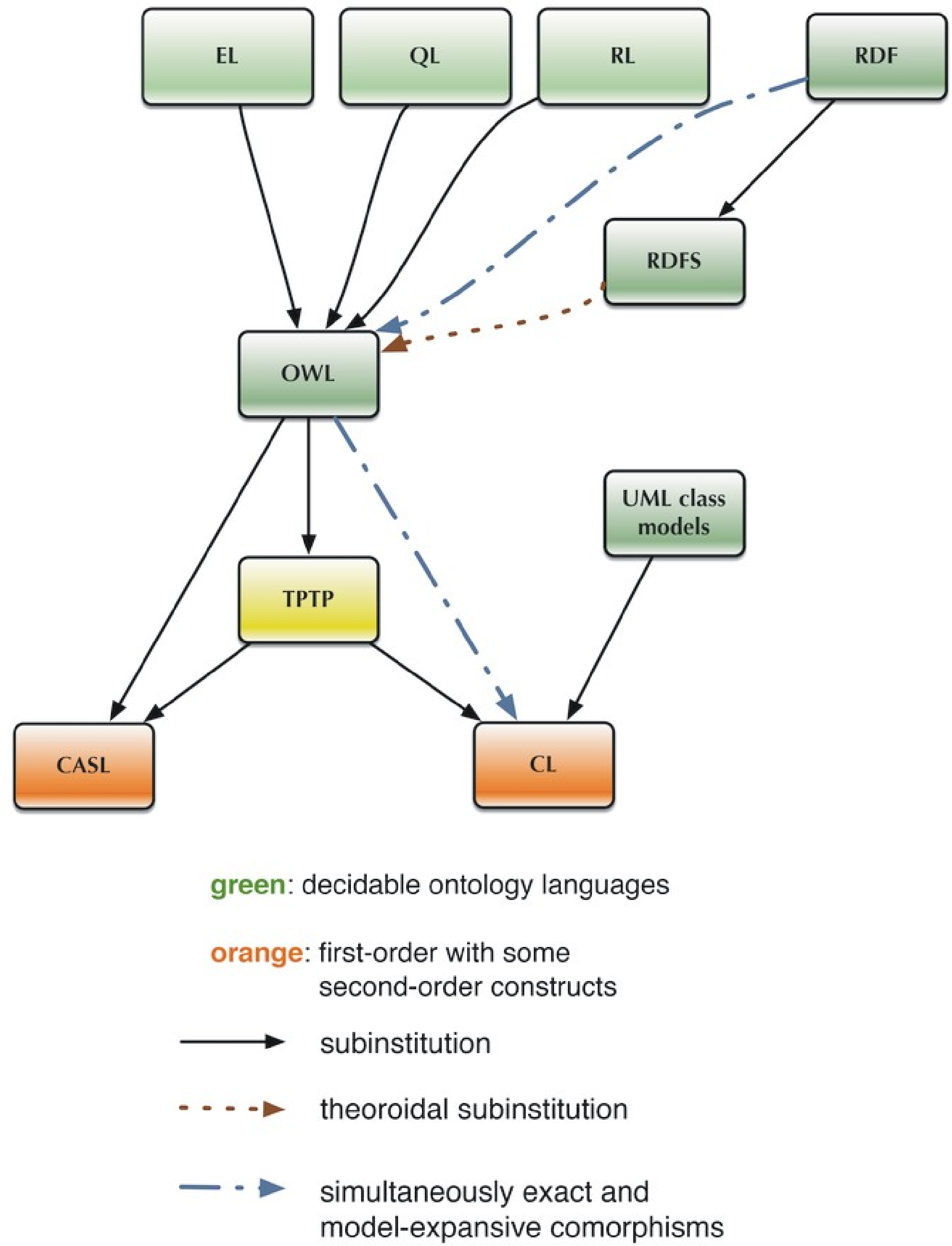
Mapping UML and the Semantic Web to CL

The diagram shows the most widely used logics supported by DOL.

Arrows show the mappings from less expressive logics to more expressive logics. Common Logic is at the lower right.

TPTP notation (for Thousands of Problems for Theorem Provers) is a version of many-sorted logic, of which classical first-order logic is a single-sorted subset.

HeTS (the Heterogeneous Tool Set) uses CASL as the interchange logic for this diagram. But other tools may use other logics.



Supporting Interoperability

A programmer's lament at a database symposium:

- *Any one of those tools, by itself, is a tremendous aid to productivity, but any two of them together will kill you.* *

Usage scenarios for DOL (Section 7 of the DOL standard):

- Interoperability between OWL and FOL ontologies
- Module extraction from large ontologies
- Interoperability between closed-world data and open-world metadata
- Verification of rules translating Dublin Core into PROV
- Maintaining different versions of an ontology in languages with different expressivity
- Metadata within OMS repositories
- Modularity and refinement of specifications
- Consistency among UML models of different types
- Refinements between UML models of different types, and their reuse
- Coherent semantics for multi-language models

*** Comment by Terry Rankin, circa 1980. But it's just as true today.**

Computability and Decidability

The logics for the Semantic Web are decidable.

- But decidabiity is a property of the problem, not the notation.
- The best TPTP systems use syntactic checks to determine the methods to use for any particular problem.
- For the same problems, those systems are as fast or faster than the tools designed for the Semantic Web.

Restricting expressive power cannot improve performance. *

- It just makes certain problems impossible to state.
- Natural languages are more expressive than any version of logic.
- But the only people who can state an undecidable sentence are those who have studied advanced logic and mathematics.
- Users always ask for more expressive power. They never ask for decidability.

* See "Fads and fallacies about logic," http://jfsowa.com/pubs/fflogic.pdf

3. Common Logic is SWeLL

A proposal for the Semantic Web Logic Language (SWeLL) evolved into the ISO/IEC standard for Common Logic (CL).*

CLIP (CL Interface to Predicate calculus) is a linear dialect of Common Logic that has a simple mapping ↔ graph logics .

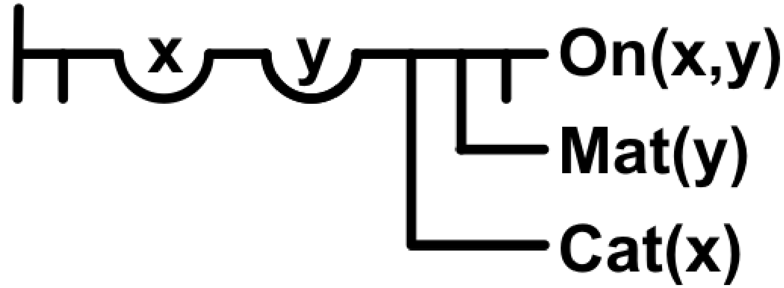
Design goals for CLIP:

- **Immediately readable by anyone who knows predicate calculus.**
- **As readable as Turtle for the RDF and OWL subsets.**
- **As readable as any notation for if-then rules.**
- **Serve as a linearization for a wide range of graph logics, including CGs, EGs, KGs, RDF, OWL, and UML diagrams.**
- **Query option: Select (*list of names*) where (*any CLIP sentence*).**
- **Support mappings of logics ↔ natural languages (NLs).**

*** The LBASE proposal: http://www.w3.org/TR/2003/NOTE-lbase-20031010**

How to say “A cat is on a mat.”

Gottlob Frege (1879):

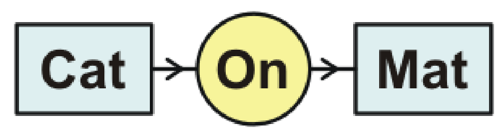


Charles Sanders Peirce (1885): $\Sigma_x \, \Sigma_y \, \text{Cat}_x \bullet \text{On}_{x,y} \bullet \text{Mat}_y$

Giuseppe Peano (1895): $\exists x \, \exists y \, \text{Cat}(x) \wedge \text{On}(x, y) \wedge \text{Mat}(y)$

Existential graph by Peirce (1897): Cat —— On —— Mat

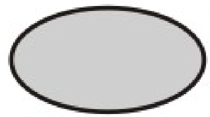
Conceptual graph (1976):



CLIP dialect of Common Logic: (∃ x y) (Cat x) (On x y) (Mat y).

Existential Graphs (EGs)

Existence: —

Negation: 

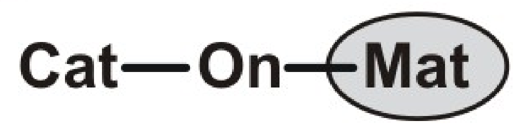
Relations: Cat• Mat• Happy• •On• •Under• •Give•

A cat is on a mat: Cat—On—Mat

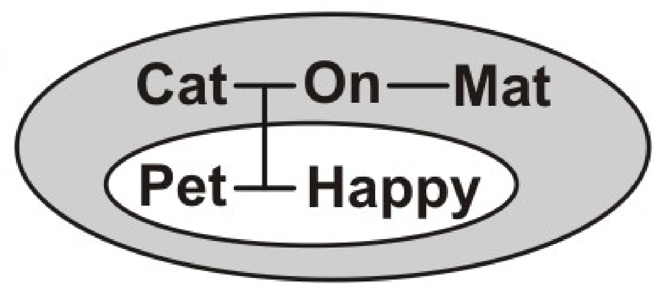
Something is under a mat: —Under—Mat

Some cat is not on a mat:

Some cat is on something that is not a mat:



If a cat is on a mat, then it is a happy pet:



The Core CLIP Notation

Existence: (∃ x) or (Exists x)

Negation: ~[ ] but ~[~[ ]] may be written [If [Then]]

Relations: (Cat x), (Mat x), (Pet x), (Happy x), (On x y), (Under x y)

***A cat is on a mat*:** (∃ x y) (Cat x) (On x y) (Mat y).

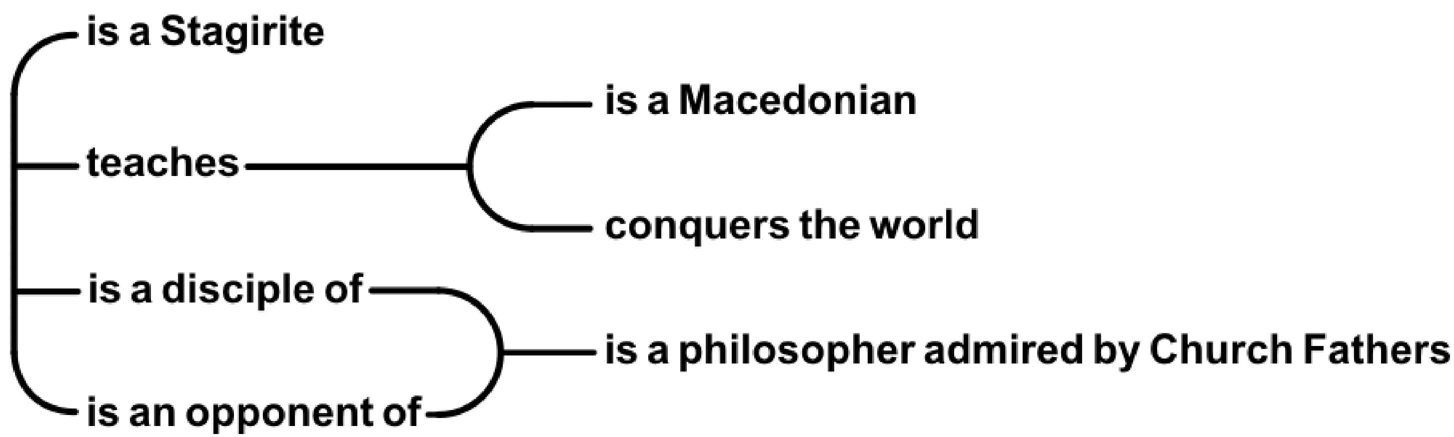
***Something is under a mat*:** (∃ x y) (Under x y) (Mat y).

***Some cat is not on a mat*:** (∃ x) (Cat x) ~[(∃ y) (On x y) (Mat y)].

***Some cat is on something that is not a mat*:**
(∃ x y) (Cat x) (On x y) ~[(Mat y)].

***If a cat is on a mat, then it is a happy pet*:**
[If (∃ x y) (Cat x) (On x y) (Mat y)
[Then (Pet x) (Happy x)]].

One of Peirce's Examples



Peirce's translation to English: ***"There is a Stagirite who teaches a Macedonian conqueror of the world and who is at once a disciple and an opponent of a philosopher admired by Fathers of the Church."***

A translation to CLIP:

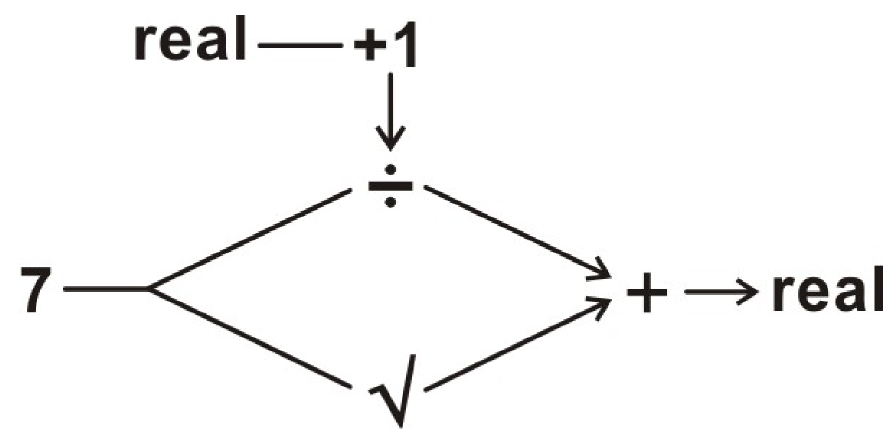
(∃ x y z) ("is a Stagirite" x) (teaches x y) ("is a Macedonian" y) ("conquers the world" y) ("is a disciple of" x z) ("is an opponent of" x z) ("is a philosopher admired by church fathers" z).

Without negation, CLIP can represent the content of a relational database, a graph database, or anything in RDF or RDFS.

Representing Functions

An example in mathematical notation: $y = 7 \div (x + 1) + \sqrt{7}$.

In EGs, a function may be represented as a relation with an arrow for its last line of identity. The four functions may be named +1, ÷, √, +.



A direct mapping of the EG to CLIP:

(∃ (x y u v w)'real) (+1 x → u) (÷ 7 u → v) (√ 7 → w) (+ v w → y).

Another option eliminates the need for the names u, v, w:

(∃ (x y)'real) (= y (+ (÷ 7 (+1 x)) (√ 7)))).

Quantifying Over Functions and Relations

CLIP allows quantified names to refer to functions and relations.

English: ***Bob and Sue are related*.**
CLIP: **(related Bob Sue).**

English: ***There is a familial relation between Bob and Sue*.**
CLIP: **(∃ r) (familial r) (relation r) (r Bob Sue).**

English: ***Every numeric function maps numbers to numbers*.**
CLIP: **(∀ f) (numeric f) (function f)**
(∀ x y) [If (f x → y) [Then (number x) (number y)]].

Literal translation of CLIP to English: ***For any numeric function f and any x and y, if f maps x to y, then x is a number and y is a number*.**

Note: Higher-order logic with a hierarchy of infinite sets is extremely inefficient for computer processing. But Common Logic uses a version of second-order logic that is as efficient as FOL.

Relating and Integrating Everything

CLIP can relate legacy systems to the latest AI tools.

- **Freely mixing and matching any notations supported by DOL.**
- **Anyone may continue to use their favorite notations indefinitely.**

Semantic Web annotations may be replaced by CLIP:

- **Any URI, enclosed in quotes, is a valid CLIP name.**
- **An annotation that uses the full expressive power of CLIP is written <clip> (*one or more CLIP sentences*) </clip>**
- **Any annotation written in a Semantic Web logic x may be rewritten <clip logic=x> (*one or more CLIP sentences*) </clip>**
- **For tools that do not support CLIP, a preprocesor may translate CLIP annotations to the corresponding SW logic.**

For integrating legacy systems with AI technology,

- **Any software that is described or specified in any UML or SW notation can take advantage of tools that process CLIP.**

4. Relating Logic to Natural Languages

For computers, informal mappings must be formalized.

- **Informal mappings to natural languages (NLs) are OK for humans.**
- **But anything a computer does is formal.**

Discourse Representation Theory specifies a subset of NLs.*

- **DRT is widely used for natural language processing (NLP).**
- **Discourse representation structures (DRSs) support full FOL.**
- **The DRS logic has a precise mapping to EG and to CLIP.**

Semi-automated translation of NLs to and from CLIP:

- **Computer translation of NL → CLIP is error prone.**
- **Computer translation of CLIP → NL is precise, but verbose.**
- **Human translation is as reliable as the human.**
- **Simpler and more reliable: Human-aided computer translation.**

*** Hans Kamp & Uwe Reyle (1993) *From Discourse to Logic*, Dordrecht: Kluwer.**

Mapping Language to Logic

Hans Kamp observed that the features of predicate calculus do not have a direct mapping to and from natural languages.

Pronouns can cross sentence boundaries, but variables cannot.

- **Example:** ***Pedro is a farmer. He owns a donkey.***
- **PC:** $\exists x\ (\text{Pedro}(x) \wedge \text{farmer}(x)).\quad \exists y\, \exists z\ (\text{owns}(y,z) \wedge \text{donkey}(z)).$
- **There is no operator that can relate x and y in different formulas.**

In English, quantifiers in the if-clause govern the then-clause.

- **Example:** ***If a farmer owns a donkey, then he beats it.***
- **But in predicate calculus, the quantifiers must be moved to the front.**
- **CLIP supports both options: English-like and PC-like.**
  If (∃ x y) (farmer x) (donkey y) (owns x y) [Then (beats x y)]].
  (∀ x y) If (farmer x) (donkey y) (owns x y) [Then (beats x y)]].

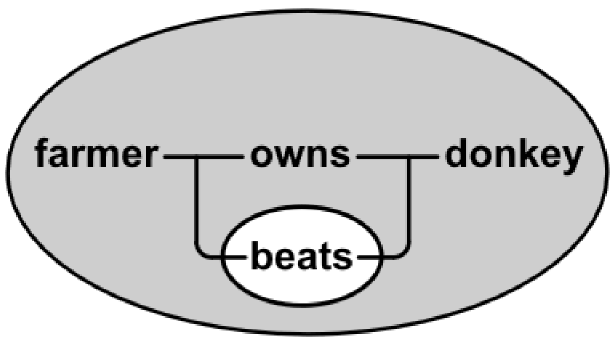
Note: Proper names are rarely unique identifiers. Both Kamp and Peirce represented names by monadic predicates.

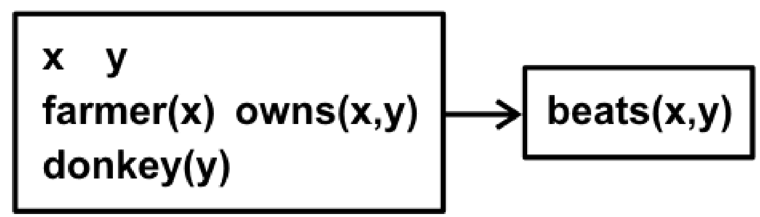
Mapping EG and DRS to CLIP

Peirce and Kamp independently chose equivalent structures.

- **Peirce chose ovals for EG with lines to show references.**
- **Kamp chose boxes for DRS with variables to show references.**
- **But the boxes and ovals represent the same logic in the same way.**

Example: ***If a farmer owns a donkey, then he beats it.***





The EG and DRS may be translated to and from exactly the same CLIP:
[If (∃ x y) (farmer x) (donkey y) (owns x y) [Then (beats x y)]].

Translating the Word *is* to Logic

Three different translations in English or CLIP:

- **Existence:** *There is x.* ↔ **(∃x)**
- **Predication:** *x is a cat.* ↔ **(Cat x)**
- **Identity:** *x is y.* ↔ **(= x y)**

Do these three translations imply that English is ambiguous?

Or is the syntax of linear notations too complex?

In EGs, all three uses of the word *is* map to a line of identity:

- **Existence:** *There is x.* ↔ —
- **Predication:** *x is a cat.* ↔ —**Cat**
- **Identity:** *x is y.* ↔ —●— **(a ligature of two lines)**

As Peirce said, EGs are more iconic than predicate calculus: they show relationships more clearly and directly.

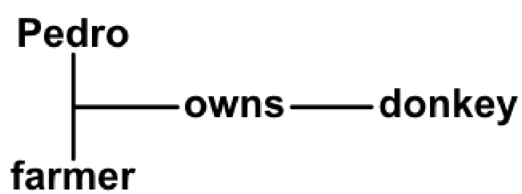
Combining EG Graphs or DRS Boxes

Two English sentences, ***Pedro is a farmer. He owns a donkey,*** **are represented by EG graphs (left) and DRS boxes (right):**


```
x
Pedro(x)  farmer(x)
```


```
y   z
owns(y,z)  donkey(z)
```

Combine them by connecting EG lines or merging DRS boxes:



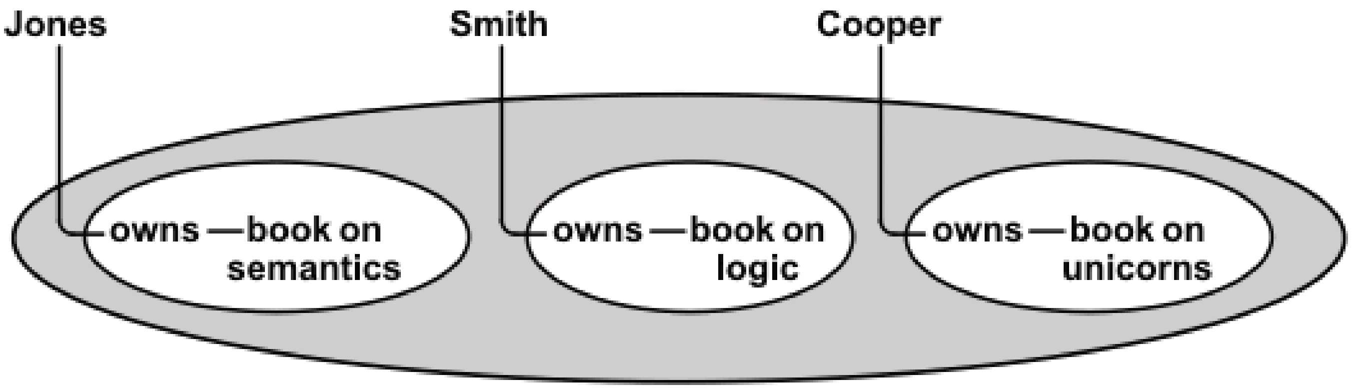

```
x   y   z   x=y
Pedro(x)  farmer(x)
owns(y,z)  donkey(z)
```

Equivalent operations on EG and DRS produce the same CLIP:
(∃ x y z) (Pedro x) (farmer x) (= x y) (owns y z) (donkey z).

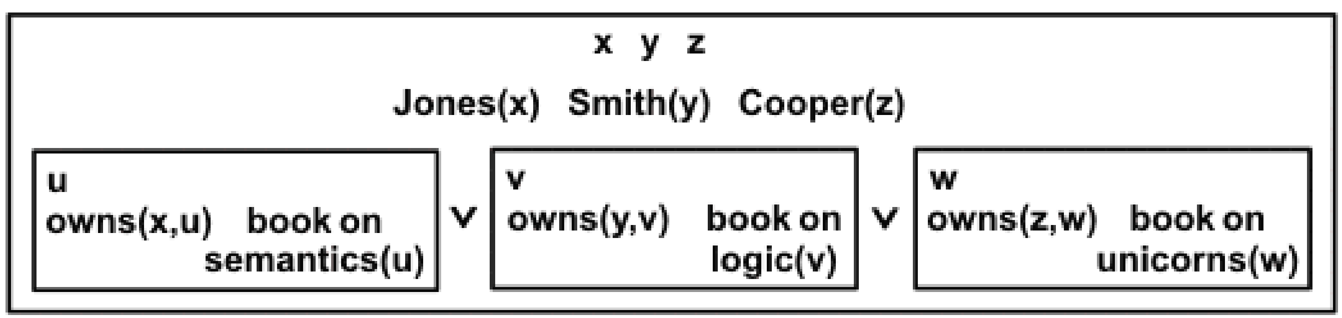
Disjunction in EG, DRS, and CLIP

Kamp and Reyle (1993): ***"Either Jones owns a book on semantics, or Smith owns a book on logic, or Cooper owns a book on unicorns."***

EG:



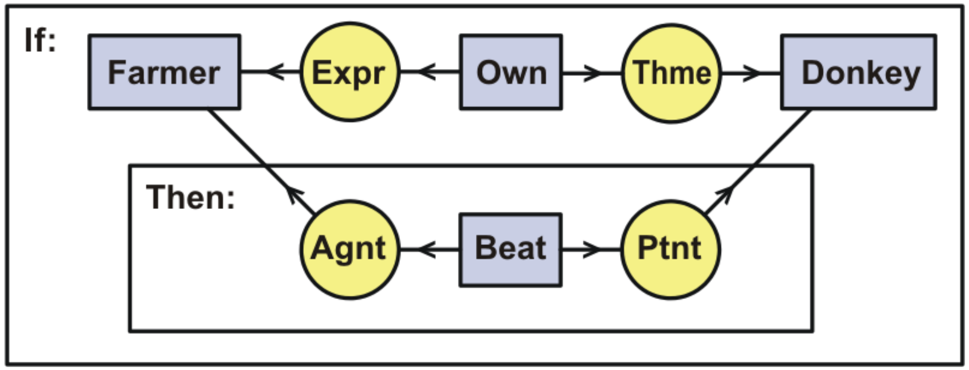
DRS:



CLIP:

(∃ x y z) (Jones x) (Smith y) (Cooper z)
[Or [(∃ u) (owns x u) ("book on semantics" u)]
[(∃ v) (owns y v) ("book on logic" v)]
[(∃ w) (owns z w) ("book on unicorns" w)]].

Conceptual Graphs



Conceptual graphs (CGs) express the same logic as EGs, but they are designed to represent the details of NL semantics. *

English: ***"If a farmer owns a donkey, then he beats it."***

CLIP: [If (∃ x'farmer y'own z'donkey) (Expr y x) (Thme y z)
[Then (∃ w'beat) (Agnt w x) (Ptnt w z)]].

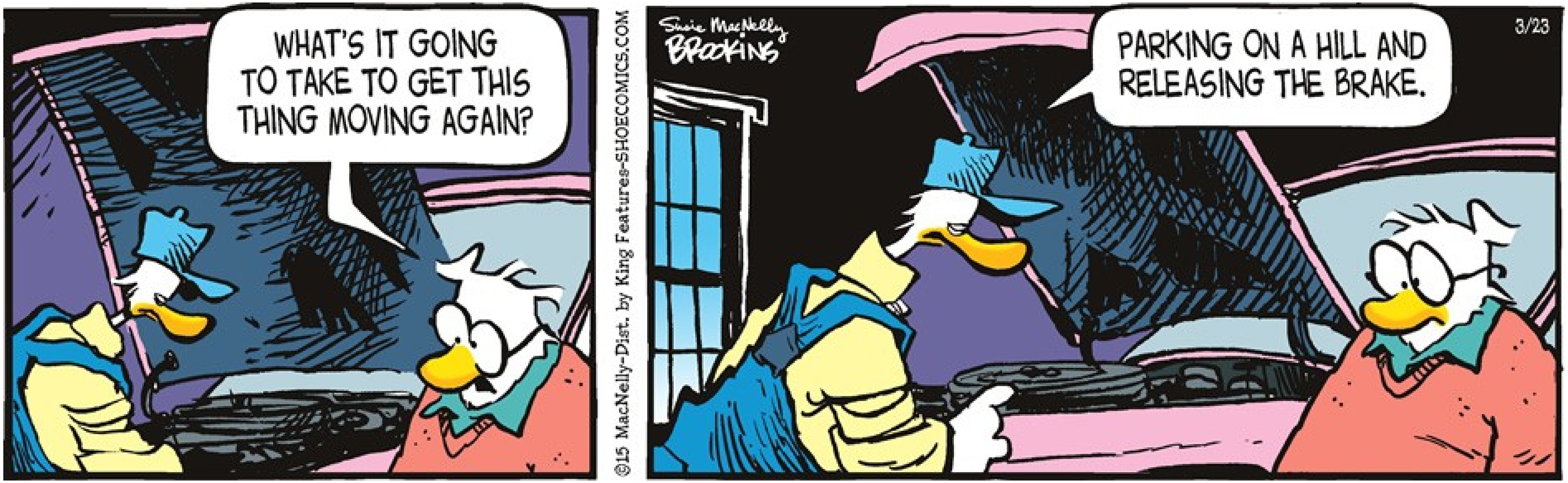
Unlike EGs, quantifiers in CGs are represented by boxes, not lines.

Names may refer to concept boxes that represent verbs.

The *semantic* or *thematic roles* used in linguistics relate verbs to nouns: experiencer (Expr), theme (Thme), agent (Agnt), and patient (Ptnt).

*** See "From EGs to CGs", http://jfsowa.com/pubs/eg2cg.pdf**

Context and Purpose



Syntax is easy: Parse the question and the answer.

Semantics depends on context and background knowledge:

- **Interpret the meaning of *thing*, *take*, and *move* In this situation.**
- **Apply the laws of physics to understand what would happen.**

Pragmatics depends on the intentions of the participants.

- **No computer system today could understand that cartoon.**
- **Computers should ask people about purpose or intentions.**

*** Source of cartoon: search for 'moving' at http://www.shoecomics.com/**

5. Metalanguage and Metadata

Metalanguage is language about language, natural or artificial.

- To define semantics, Tarski (1933) used logic as a metalanguage for defining the truth value of any statement in logic.
- Annotations in SW logics state metadata about documents.
- But metadata about the metadata can also be useful.

The IKL extension to Common Logic supports metalanguage. *

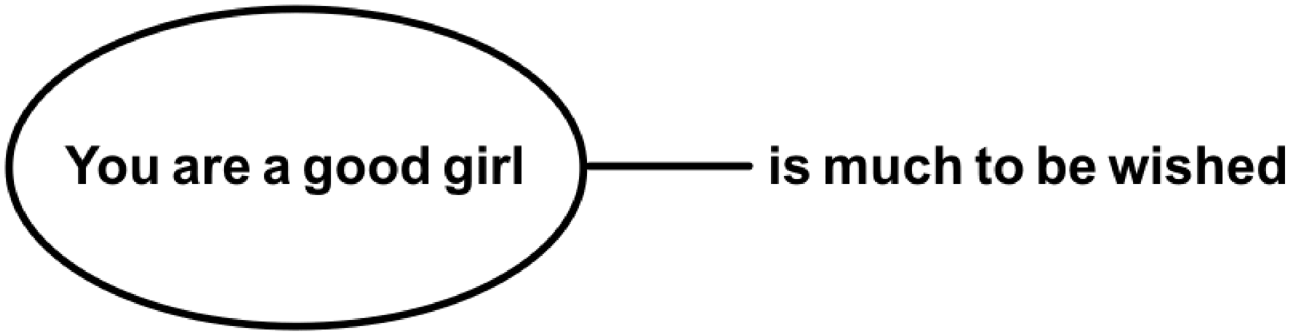
- IKL enables CLIP to comment on anything expressed in CLIP.
- It can represent metadata about the sources and reliability of data.
- It can support reasoning about metaphor, modality, and the issues of vague, fuzzy, missing, erroneous, or fraudulent information.

Any Unicode strings may be used for CLIP names. Metadata may even be expressed by emojis.

* For the IKL documents, see http://jfsowa.com/ikl/ .

Metalanguage in Existential Graphs

A metalevel EG by Peirce (1898):



Peirce's English: "*That you are a good girl is much to be wished.*"

A shaded oval negates the nested EG. Without shading, the EG expresses a proposition that is neither asserted nor negated.

The same proposition in CLIP: ["You are a good girl"].

History: *From 1898 to 1914, Peirce wrote extensively about metalanguage, modality, and intentionality. Those writings had a strong influence on logicians, philosophers, linguists, and AI researchers. For references, see* **http://jfsowa.com/pubs/5qelogic.pdf**

Metalanguage About Situations

The drawing on the right may be interpreted in three ways.

1. Actual: *Pierre is thinking of Marie, who is thinking of him.*

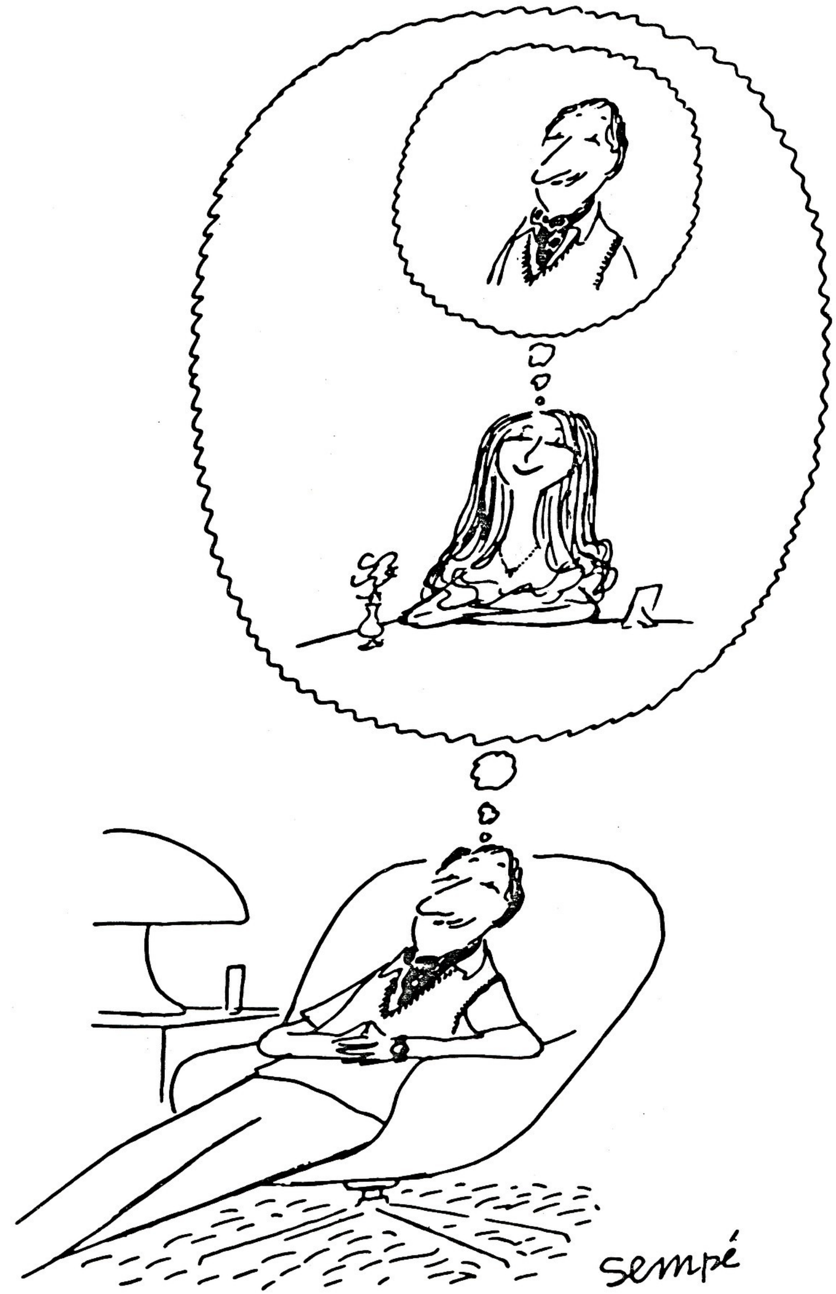
2. Modal: *Pierre is thinking of Marie, who may be thinking of him.*

3. Intentional: *Pierre hopes that Marie is thinking of him.*

In the second clause of #1, the verb *is* implies that Pierre's thought is true.

In #2, the verb *may* implies that his thought is a possible proposition.

In #3, the object of the verb *hopes* is a situation Pierre intends in some way.



Metalanguage in CLIP

Peirce's example of 1898 represents an intended situation.

English: ***That you are a good girl is much to be wished*** **[*****by someone*****].**

CLIP: ("is much to be wished" [Situation "You are a good girl"]).

English and CLIP for the sentences about Pierre.

English: ***Pierre is thinking of Marie, who is thinking of him.***

CLIP: (thinkingOf Pierre Marie) (thinkingOf Marie Pierre).

English: ***Pierre is thinking of Marie, who may be thinking of him*****.**

CLIP: (thinkingOf Pierre Marie) (possible [(thinkingOf Marie Pierre)]).

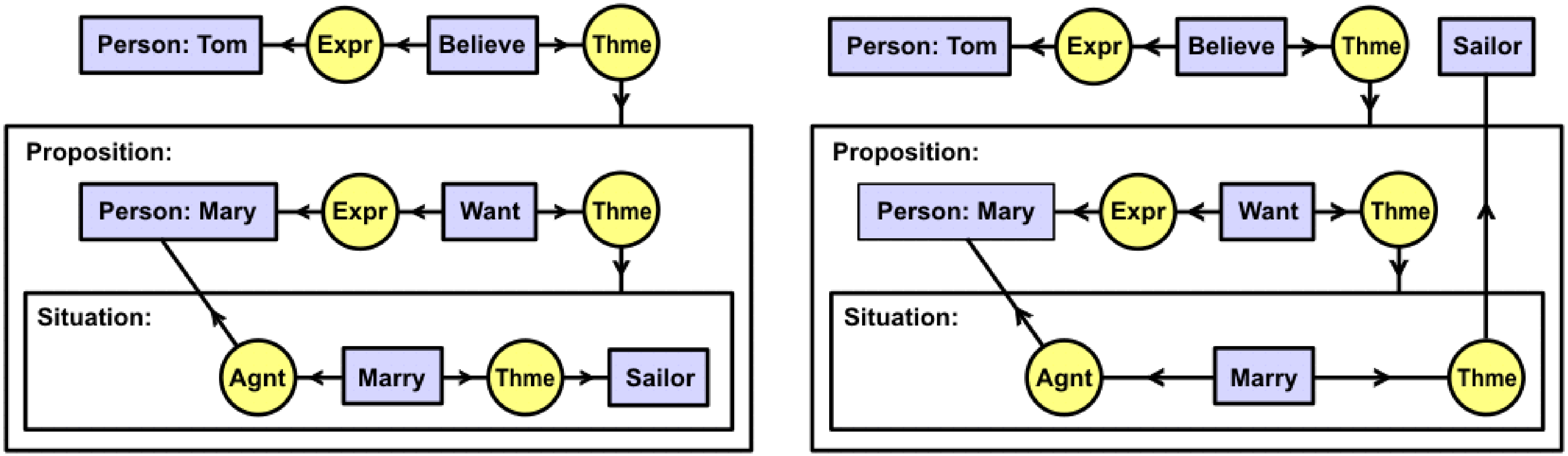
English: ***Pierre hopes that Marie is thinking of him*****.**

CLIP: (hopesFor Pierre [Situation (thinkingOf Marie Pierre)]).

By itself, IKL does not support modal logic. But IKL at the metalevel can be used to define modal relations in terms of laws and facts.*

*** See http://jfsowa.com/pubs/worlds.pdf**

CGs for Propositions and Situations



The CGs above show two of the three interpretations of the sentence ***Tom believes that Mary wants to marry a sailor:***

- ***Tom believes a proposition that Mary wants a situation in which there exists a sailor whom she marries.***
- ***There is a sailor, and Tom believes that Mary wants to marry him.***

For the third interpretation, the blue box for Sailor would be moved to the area of the proposition:

- ***Tom believes that there is a sailor whom Mary wants to marry.***

6. From Perception to Cognition

Minsky: Cognition as a diversity of interacting processes.

"What magical trick makes us intelligent? The trick is that there is no trick. The power of intelligence stems from our vast diversity, not from any single, perfect principle. Our species has evolved many effective although imperfect methods, and each of us individually develops more on our own. Eventually, very few of our actions and decisions come to depend on any single mechanism. Instead, they emerge from conflicts and negotiations among societies of processes that constantly challenge one another." *

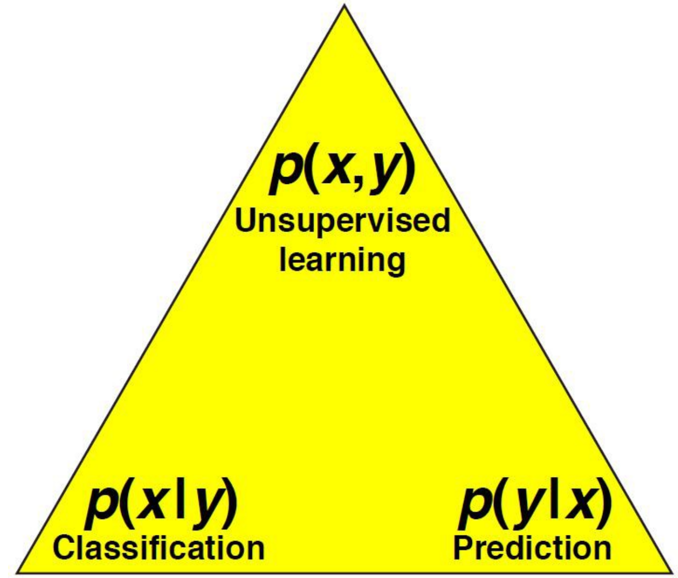
Barsalou: Cognition as a coordinated system of systems.

Cognition "emerges from deep dependencies between all the basic systems in the brain, including goal management, perception, action, memory, reward, affect, and learning. We also believe that human cognition greatly reflects its social evolution and context." **

* Marvin Minsky (1986) *The Society of Mind*, New York: Simon & Schuster.

** L. W. Barsalou, C. Breazeal, & L. B. Smith (2007) Cognition as coordinated non-cognition, http://barsaloulab.org/Online_Articles/pdf

Machine Learning (ML)



Most ML methods learn to approximate a function f: $x \rightarrow y$, where *x* and *y* are vectors of observable features or attributes. *

Unsupervised learning begins with a set of pairs of the form (*x,y*) and computes an estimated probability *p*(*x,y*) for any *x* and *y*.

For classification, *p*(*x* | *y*) is the probability that something described by the feature vector *y* belongs to a class *x*.

For prediction, *p*(*y* | *x*) is the probability that a state described by a vector of features *x* will be followed by a state described by *y*.

Such functions represent the kind of learning that psychologists analyzed and described by stimulus-response (S-R) theories.

But S-R theories could not explain novelty, discovery, intentionality, or the open-ended diversity analyzed by Minsky and Barsalou.

* Henry Lin & Max Tegmark (2016) Why does deep and cheap learning work so well?

Applications of Machine Learning *

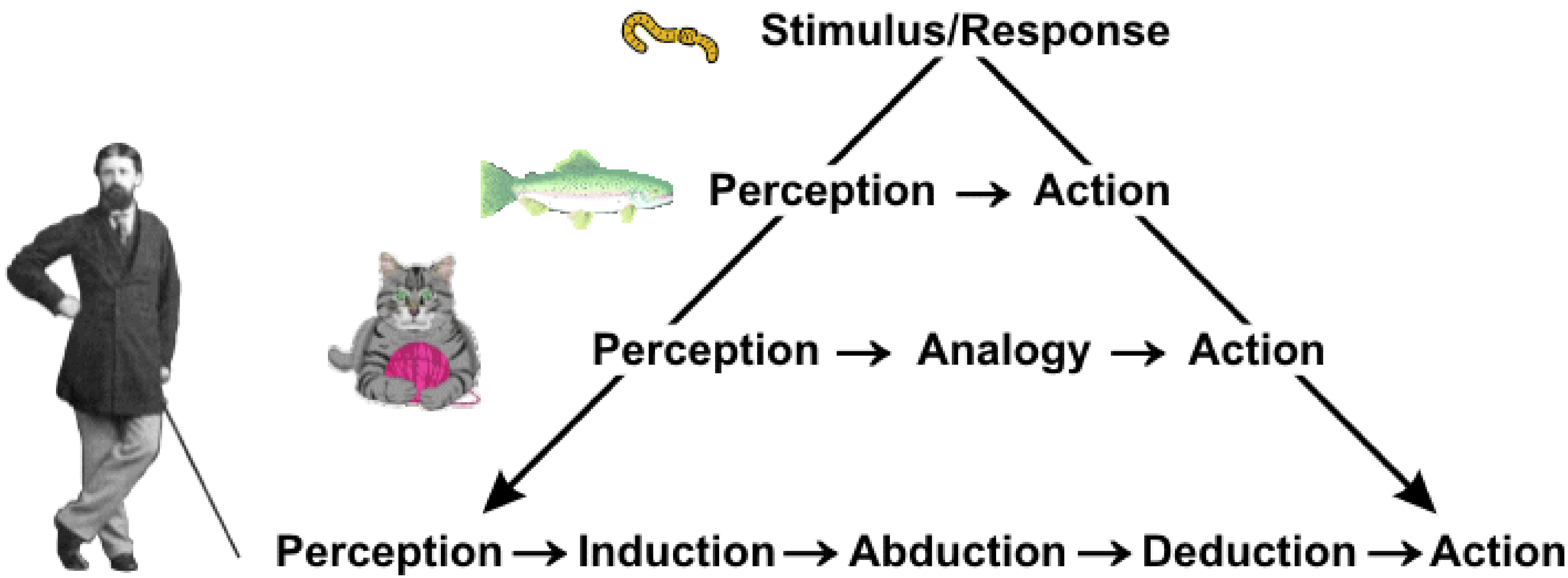
Stimulus x	Response y	Application
Picture	Are there human faces? (0 or 1)	Photo tagging
Loan application	Will they repay the loan? (0 or 1)	Loan approvals
Ad plus user information	Will user click on ad? (0 or 1)	Targeted online ads
Audio clip	Transcript of audio clip	Speech recognition
English sentence	French sentence	Language translation
Sensors from hard disk, plane engine, etc.	Is it about to fail?	Preventive maintenance
Car camera and other sensors	Position of other cars	Self-driving cars

Learning a function f: $x \rightarrow y$ is the basis for perception.

- **Observation by Andrew Ng: Current ML methods automate tasks that take less than one second of mental effort by humans.**
- **Every one of Ng's examples recognizes a pattern.**
- **None of them do complex reasoning or language understanding.**

*** Andrew Ng (2016) https://hbr.org/2016/11/what-artificial-intelligence-can-and-cant-do-right-now**

Thinking Beyond the First Second



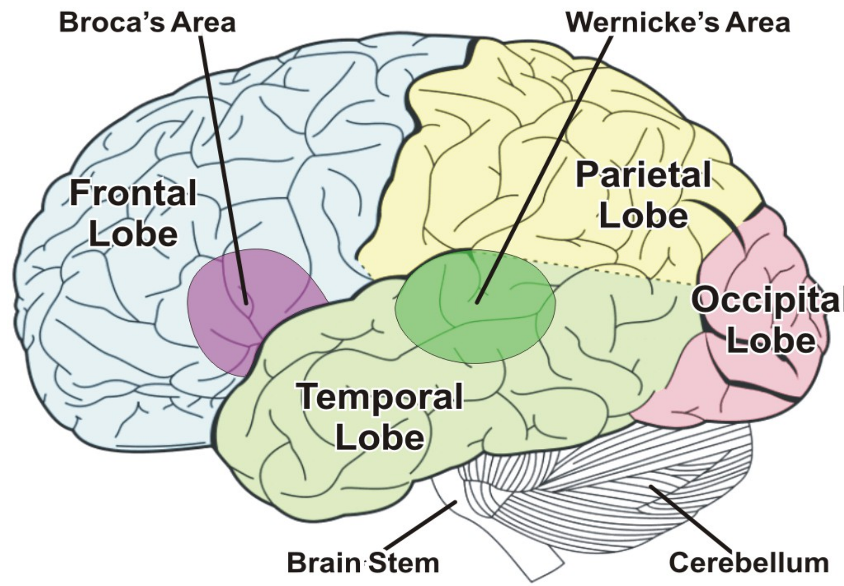
Perception and classification take one second or less.

- **Neural nets are valuable for learning and recognizing patterns.**
- **By themselves, NNs support a fish level of intelligence.**
- **With analogies, NNs can support a cat level of intelligence.**

Analysis, planning, discovery, and innovation take more time.

- **They require cycles of induction, abduction, deduction...**

Cognitive Learning



The areas of the cerebral cortex are highly specialized.

A study with fMRI scans showed which areas are active at various stages of learning. *

14 participants studied how four devices work: bathroom scale, fire extinguisher, automobile braking system, and trumpet.

Cognitive learning is much deeper than deep neural nets:

1. Occipital lobes are active in recognizing shapes and patterns.
2. Parietal lobes become active in learning mechanical structures.
3. All lobes become active as participants are "generating causal hypotheses" about how the system works.
4. Finally, the frontal lobes anticipate "how a person (probably oneself) would interact with the system."

* R. A. Mason & M. A. Just (2015) http://medicalxpress.com/news/2015-03-science-brain.html

Areas Active in Cognition

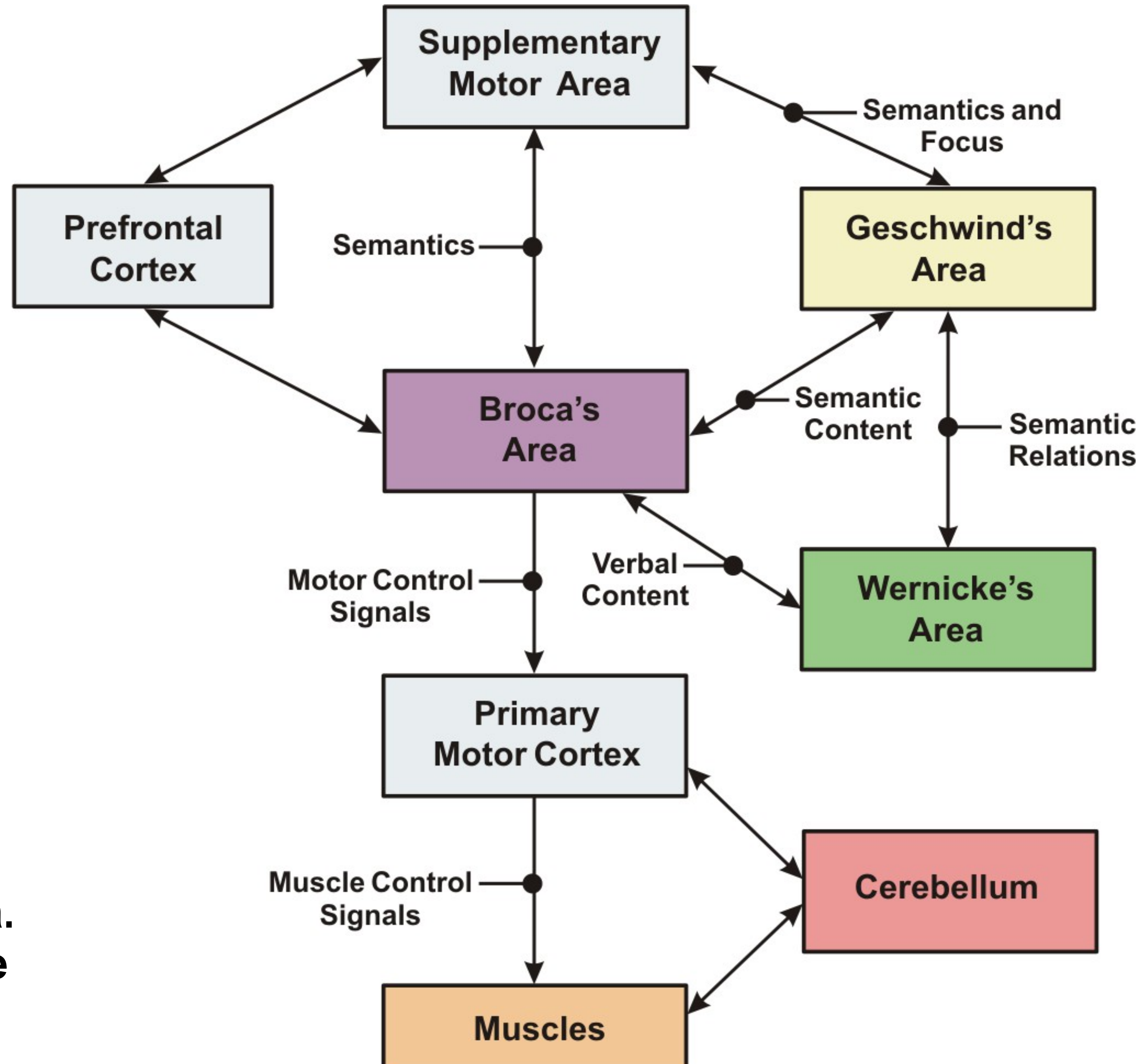
Most neurons have short links to nearby neurons.

But others make long-distance connections from one lobe to another.

The diagram shows connections among areas of the brain involved in language. *

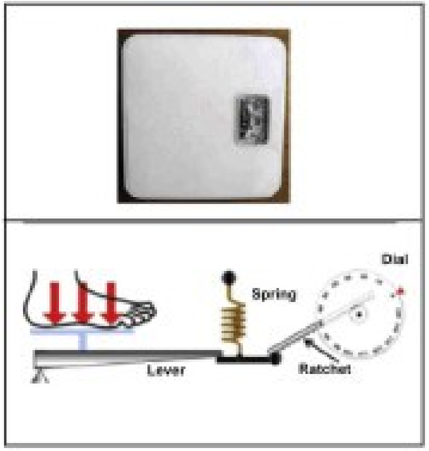
The colors of the boxes correspond to the colors of the brain areas in the previous slide.

Patterns can be learned and recognized in one area. But cognition links diverse areas across the brain.



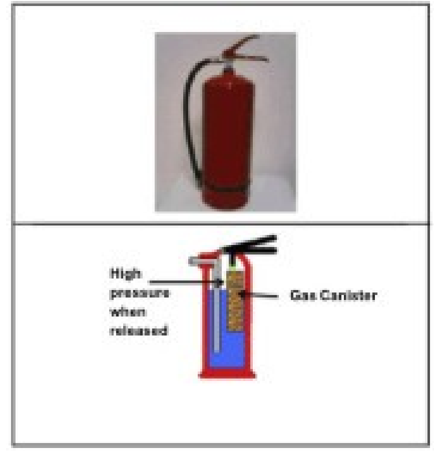
* Diagram adapted from MacNeilage (2008).

Perceptual and Cognitive Learning

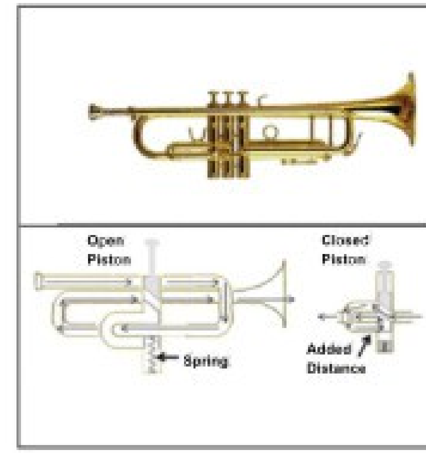


Bathroom Scale

Disc Brake System



Fire Extinguisher



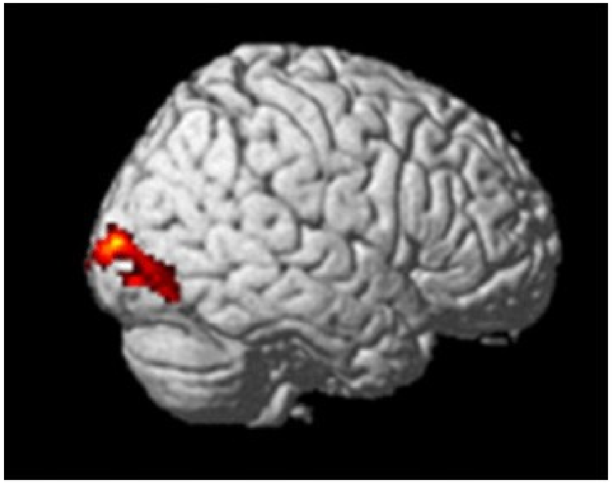
Trumpet

14 participants studied how four devices work: bathroom scale, fire extinguisher, disc brake system, and trumpet. *

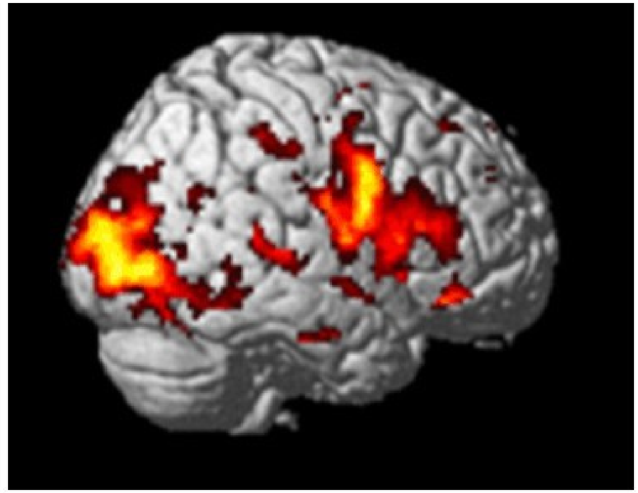
- **Subjects: college students who were not science or engineering majors.**
- **They had multiple training sessions with each of the four devices.**
- **During test sessions, an fMRI scanner recorded patterns of brain activity.**
- **An early training session just showed pictures and named the parts:** ***A bathroom scale consists of a spring, a lever, a ratchet, and a dial.***
- **Later sessions explained structural and causal relations:** ***The spring pulls a ratchet which rotates a gear attached to a measurement dial.***

*** R. A. Mason & M. A. Just (2015) Physics instruction induces changes in neural representation.**

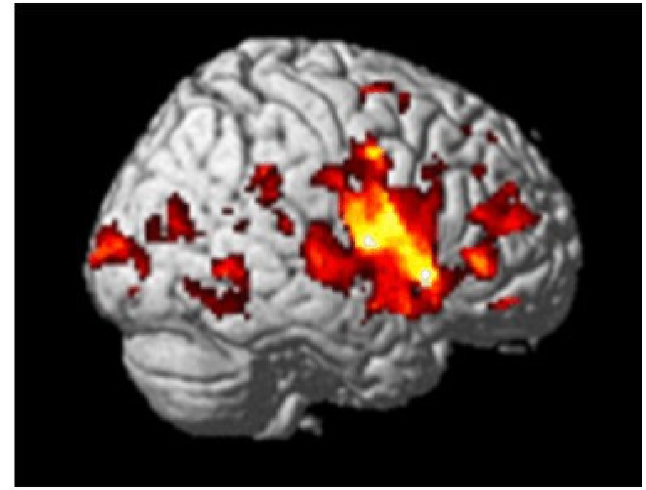
Cognitive Learning



1. Visual perception



2. Thinking about structure



3. Thinking about causality

Neural activity in the right hemisphere during test sessions:

- **All 14 students showed similar neural activations.**
- **Questions about the objects and parts activated the visual cortex, the occipital lobes in the back of the brain (fMRI image #1 above).**
- **Questions about structural relations activated the parietal lobes, which link vision to all sensory and motor regions (image #2).**
- **Questions about the causal effects of someone operating the system activated the frontal lobes and connections across the brain (image #3).**
- **Summary: Cognitive learning involves structural and causal relations that link and coordinate perception, action, and reasoning.**

Peirce's Trichotomies

While studying Kant, Peirce analyzed the patterns of triads in Kant's table of 12 categories (4 x 3).

He discovered metalevel patterns underlying those categories:

- First: Quality expressible by a monadic predicate.
- Second: Reaction expressible by a dyadic relation.
- Third: Mediation that relates a first and a second.

Basic trichotomy: Some observable Mark (1) may be interpreted as a Token (2) of some Type (3). *

- The phaneron: "whatever is throughout its entirety open to direct observation" (Peirce, MS 337, 1904).
- It's a continuum of possible marks prior to any interpretation.
- Peirce compared phaneroscopy to the work of artists who can draw marks as seen before any interpretation.
- That talent enables artists to imagine what people will see and feel as they walk through buildings that have not yet been built.

* See Signs and Reality (Sowa 2015) for examples and discussion of Peirce's categories.

Intentionality

Without life, there is no meaning in the universe.

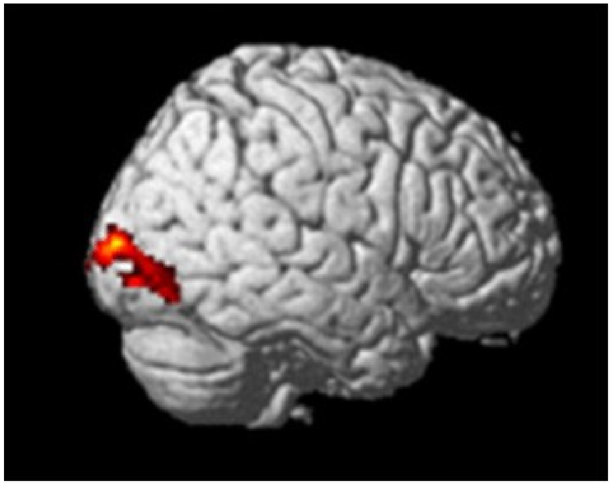
- **Philosopher Franz Brentano: Intentionality is "*the directedness of thought toward some object, real or imagined.*"**
- **Biologist Lynn Margulis: "*The growth, reproduction, and communication of these moving, alliance-forming bacteria become isomorphic with our thought, with our happiness, our sensitivities and stimulations.*" ***
- **A bacterium swimming upstream in a glucose gradient marks the beginning of goal-directed intentionality.**

In Peirce's categories, intentionality is a mediating Third.

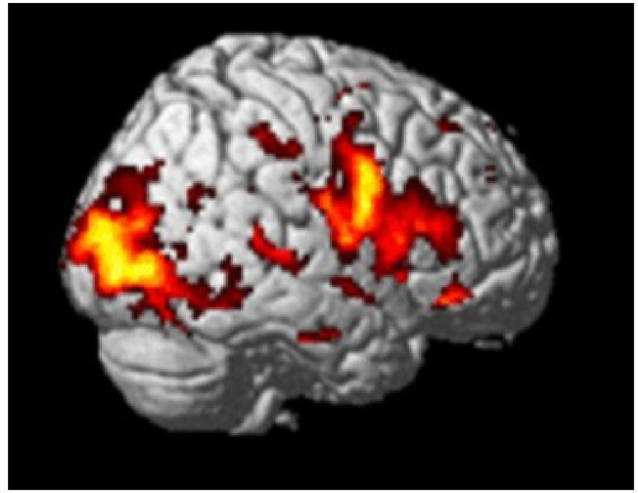
- **It's the reason why some mind or quasi-mind directs attention toward some mark, which it interprets as a token of some type.**
- **Some interpretation by some agent makes some mark (an aspect of the universe) meaningful in some way for that agent.**
- **All laws, communications, explanations, value judgments, and social relations depend on the intentions of some agent.**

*** Margulis (1995) Gaia is a tough bitch, http://edge.org/documents/ThirdCulture/n-Ch.7.html**

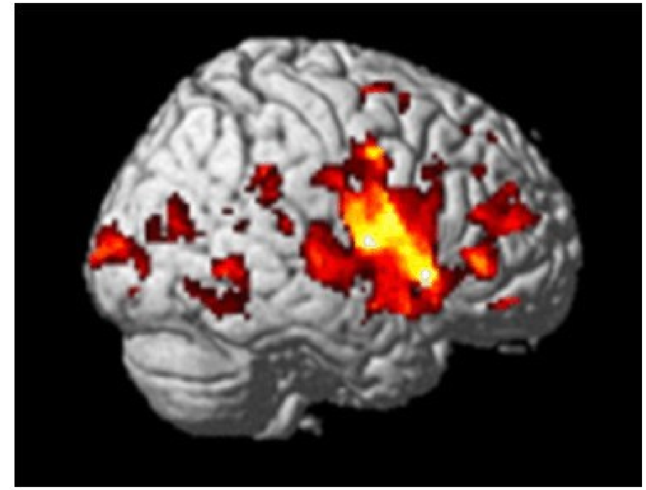
Neuro-Symbolic Reasoning



1. Visual perception



2. Thinking about structure



3. Thinking about causality

Neural correlates of Peirce's First, Second, and Third:

- Perception is based on localized percepts or prototypes. It classifies phenomena by monadic predicates (fMRI image #1).
- Long-distance connections in the parietal lobes support dyadic relations that connect all sensory and motor modalities (image #2).
- The frontal lobes process the mediating triadic relations in reasoning, planning, causality, and intentionality (image #3).
- Much more detail must be analyzed and explained, but these examples suggest promising directions for future R & D.

Cognitive Memory (CM)

CM is an associative memory for reasoning by analogy. *

- **Associative storage and retrieval of graphs in log(N) time.**
- **Precise pattern matching (unification) for logic and mathematics.**
- **Approximate pattern matching for analogies and metaphors.**

Formal reasoning is based on a disciplined use of analogy:

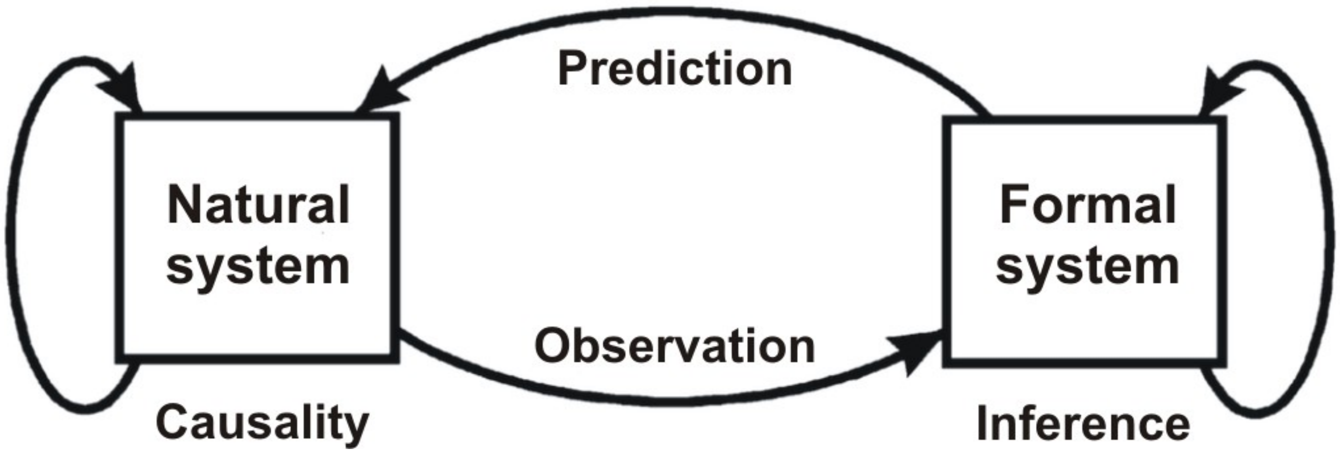
- **Induction: Generalize multiple cases to create rules or axioms.**
- **Deduction: Match (unify) part of a new case with some rule or axiom.**
- **Abduction: Form a hypothesis based on aspects of similar cases.**

Informal analogies support more general cognitive reasoning:

- **CM can store large volumes of previous knowledge and experience.**
- **Any new case can be matched to similar cases in long-term memory.**
- **Close matches are ranked by a measure of semantic distance.**

*** Survey of CM and applications: http://jfsowa.com/talks/cogmem.pdf**

Anticipatory Systems



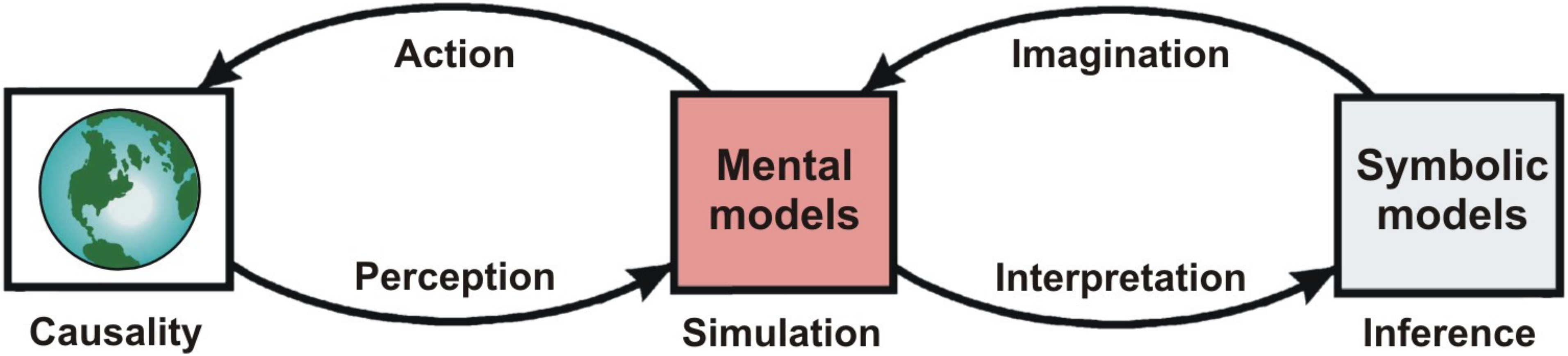
Rosen (1985) drew this diagram to explain anticipation.

- **People and other animals make inferences from observations.**
- **Inference from past events can predict future possibilities.**
- **Therefore, anticipation is caused by the past, not the future..**

But animals and young children don’t study logic.

- **Their inferences cannot be based on a system of formal logic.**
- **Even adults perform complex tasks without using formal logic.**
- **What kind of reasoning do they use?**

Neuro-Symbolic Hybrid



The human brain can support two kinds of reasoning:

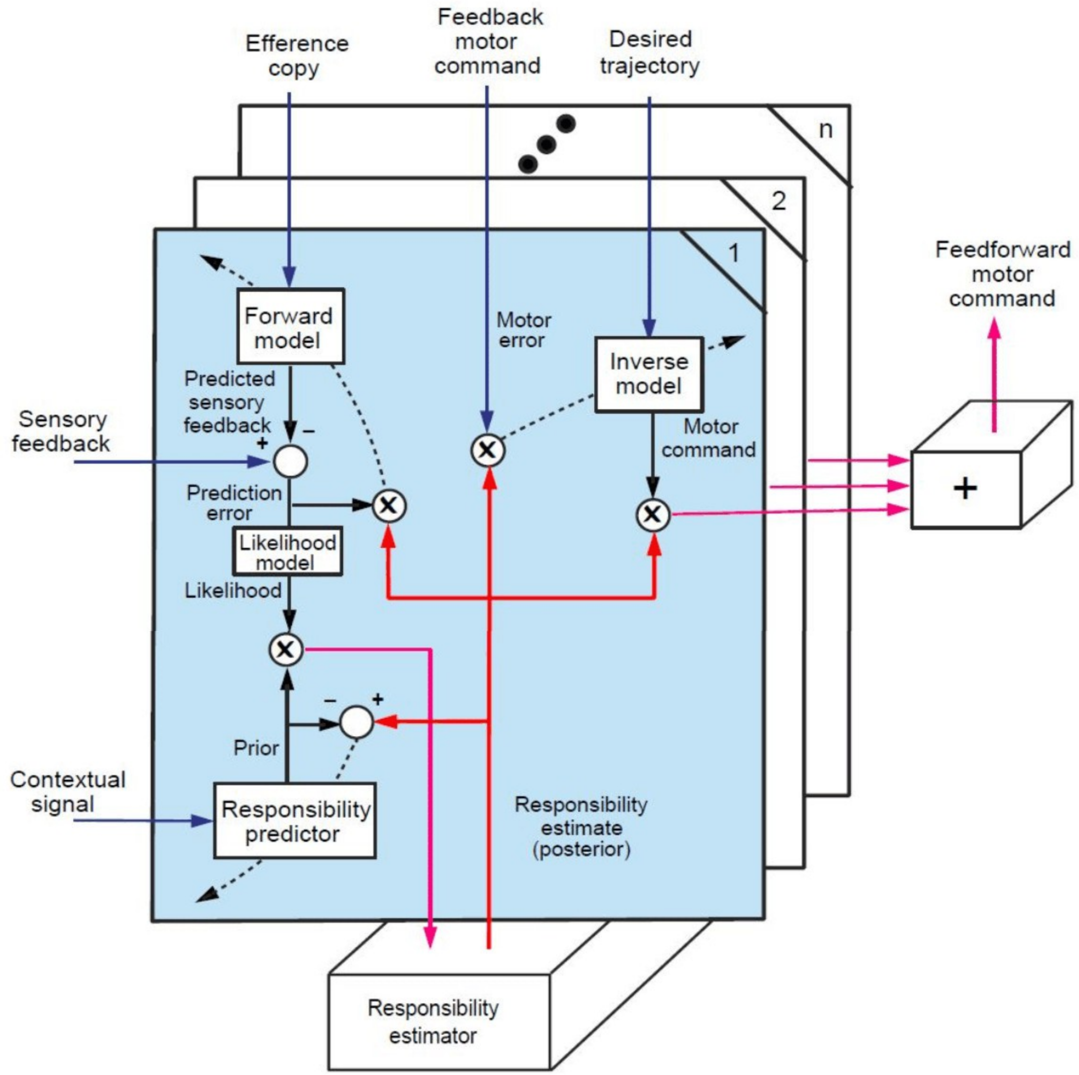
- Symbolic reasoning is necessary for language and logic.
- But symbols have an indirect mapping to and from the world.
- Mental models are fundamental for humans and other animals.
- Action and perception relate mental models to the world.

Neuro-symbolic hybrids combine the advantages of both.

- Mental models are simulated in the cerebellum. Inferences about symbolic models are processed in the frontal lobes. The two are complementary. *

* See "Cortical models," http://dx.doi.org/10.1098/rstb.2016.0515
See also "Two paradigms are better than one," http://jfsowa.com/talks/paradigm.pdf

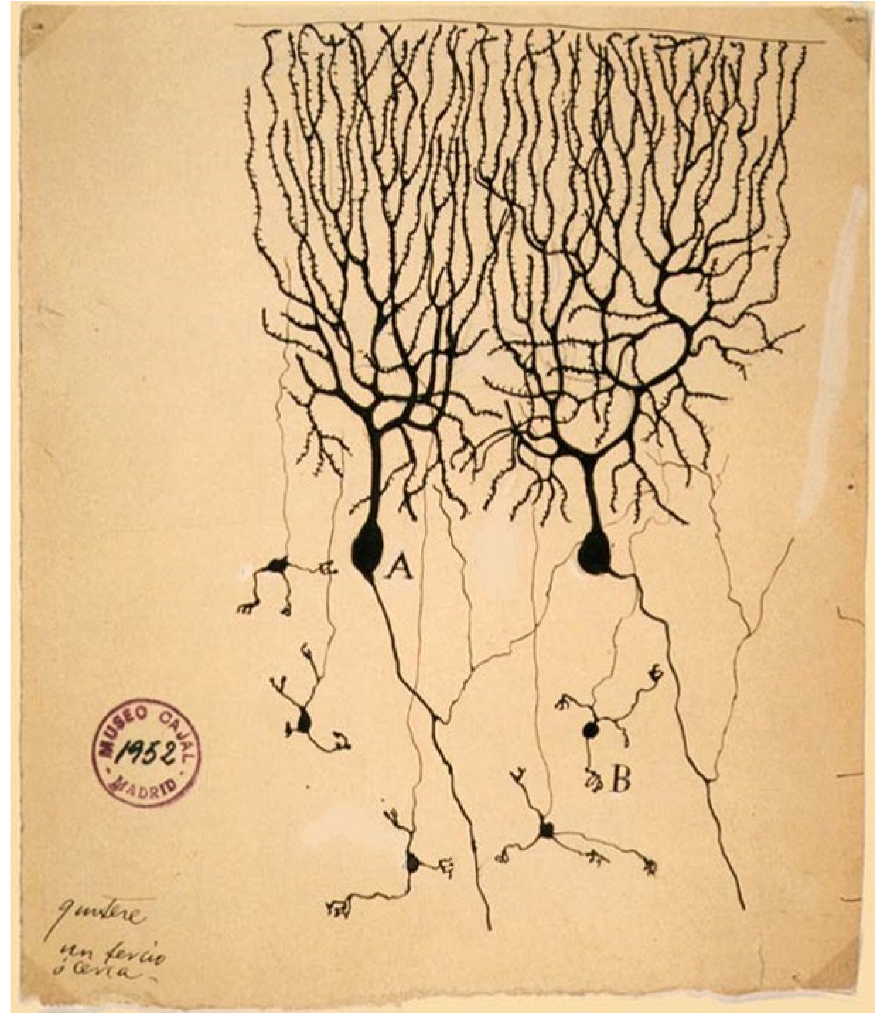
Reasoning by the Cerebellum



The cerebellum can make predictions (forward reasoning) or determine responsibility (backward reasoning). It relates perception and intention. *

*** Diagram from Wolpert, Miall, & Kawato (1998).**
See also Roth, Synofzik, & Lindner (2013) The Cerebellum optimizes perceptual predictions .

Computational Power



Drawing shows Purkinje neurons (A) and granule neurons (B). *

Cerebellum has about 50 billion granule cells, each with about 4 dendrites.

It has much fewer Purkinje cells, but each one has over 100,000 synapses.

Marr, Albus, and others suggested that granule cells form networks that learn perceptual and motor patterns.

Purkinje cells transmit that information from the granule cells to the cerebrum.

The cerebrum builds on the perceptual patterns as the basis for cognitive learning, reasoning, and language.

The cerebrum is a general-purpose processor, and the cerebellum behaves like a computationally intensive graphics processor.

*** Drawing by Santiago Ramón y Cajal, copy from Wikipedia.**

The Role of the Cerebellum

Different activities require different computational power.

- **Fruit bats eat stationary food in daylight hours.**
- **Insect-eating bats catch flying food in the dark.**

The role of the cerebellum in bats, birds, and dolphins:

- **The lobes of the cerebellum of both kinds of bats have similarities to the lobes of birds, because they fly in 3-dimensional space.**
- **The lobes of the cerebellum of dolphins also have similarities to the lobes of birds, because dolphins swim in 3-D space.**
- **The auditory lobes of the cerebellum for dolphins and the insect-eating bats have similarities because they both use echolocation.**
- **The greater sensory complexity for insect-eating bats is correlated with a relatively larger cerebellum.**

The cerebellum simulates a virtual reality of the mind.

- **It processes 3-D sensory-motor data for perception and action.**
- **It has direct connections to and from the occipital lobes.**

Mathematics and the Cerebellum

The cerebellum is active in mathematical reasoning.

- Neuroscientists had thought it was dedicated to controlling movement.
- But fMRI scans show that it's also active when people are doing math.

It's even active for the most abstract branches of mathematics:

- Early studies showed that the cerebellum is active in arithmetic. *
- But it's active in expert reasoning in every branch of mathematics.

Comparison of mathematicians and non-mathematicians: **

- Subjects: Experts in algebra, analysis, geometry, or topology.
- Controls: Same academic standing, but mathematically naive.
- During an fMRI scan, they classified sentences as meaningful or meaningless. If meaningful, they also responded true or false.
- On sentences about math, the cerebellum was active for mathematicians, but not for the non-mathematicians.
- This experiment confirms introspective reports by mathematicians.

* S. Feng et al. (2008) The cerebellum connectivity in mathematics cognition.
** M. Almaric & S. Dehaene (2016) Brain networks for advanced mathematics.

Mathematics and Imagery

Paul Halmos, mathematician:

“Mathematics — this may surprise or shock some — is never deductive in its creation. The mathematician at work makes vague guesses, visualizes broad generalizations, and jumps to unwarranted conclusions. He arranges and rearranges his ideas, and becomes convinced of their truth long before he can write down a logical proof... the deductive stage, writing the results down, and writing its rigorous proof are relatively trivial once the real insight arrives; it is more the draftsman’s work not the architect’s.” *

Albert Einstein, physicist:

“The words or the language, as they are written or spoken, do not seem to play any role in my mechanism of thought. The psychical entities which seem to serve as elements in thought are certain signs and more or less clear images which can be *voluntarily* reproduced and combined... The above-mentioned elements are, in my case, of visual and some of muscular type. Conventional words or other signs have to be sought for laboriously only in a secondary stage, when the mentioned associative play is sufficiently established and can be reproduced at will.” *

*** For details and sources, see http://jfsowa.com/talks/ppe.pdf**

Future Directions for AI

Neuro-symbolic integration can make AI more human-like.

- Neural nets are necessary for perception and recognition.
- Symbolic methods are necessary for language and logic.
- High-speed graphics are necessary for virtual reality.

To see the need for all three, consider sharks and dolphins:

- They are about the same size, and they hunt the same prey.
- A shark has a large cerebellum, but a tiny forebrain.
- A dolphin has a huge cerebellum and a huge cerebral cortex.
- A group of sharks devour their prey in a food frenzy.
- But a group of dolphins can communicate, organize a hunt, and systematically surround and trap their prey.
- Dolphins also train their young, care for each other, and are friendly with humans — desirable traits in AI systems.

*See Neural-symbolic integration and the Semantic Web.
See also Neuro-symbolic AI is the future of artificial intelligence.

7. Automated and Semi-automated Tools

The tools should support a dialogue.

- Explanation requires more interaction than question-answering.
- Both novices and experts should be able to carry on an open-ended conversation about any subject they choose.
- Follow-up questions may drill down to any depth required.
- Computers should accept any language or notation people prefer, and they should read documents without requiring prior annotations.
- If a computer can’t understand some text, it should ask people for help. People should answer in their own language.
- Computers may annotate texts, but they may need human assistance.

A dialogue should be as precise or vague as the subject matter.

- Human languages can describe a continuous, dynamically changing world at any level of detail and precision.
- A dialogue with computers should be just as flexible.
- Requirement: Neuro-symbolic learning and reasoning.

Advice From Two Logicians and a Poet

Alfred North Whitehead:

"Human knowledge is a process of approximation. In the focus of experience, there is comparative clarity. But the discrimination of this clarity leads into the penumbral background. There are always questions left over. The problem is to discriminate exactly what we know vaguely."

Charles Sanders Peirce:

"It is easy to speak with precision upon a general theme. Only, one must commonly surrender all ambition to be certain. It is equally easy to be certain. One has only to be sufficiently vague. It is not so difficult to be pretty precise and fairly certain at once about a very narrow subject."

Alfred North Whitehead:

"We must be systematic, but we should keep our systems open."

Robert Frost:

"I've often said that every poem solves something for me in life. I go so far as to say that every poem is a momentary stay against the confusion of the world... We rise out of disorder into order. And the poems I make are little bits of order."

[To make the comparison, replace every occurrence of *poem* with *theory*.]

Formal Concept Analysis (FCA)

A theory and tools for semi-automated ontology design:

- Theory. Define a minimal lattice that shows all inheritance paths among a set of concept types, each defined by a list of attributes.
- Algorithms. Efficient ways for computing a minimal lattice from a list of terms and defining features.

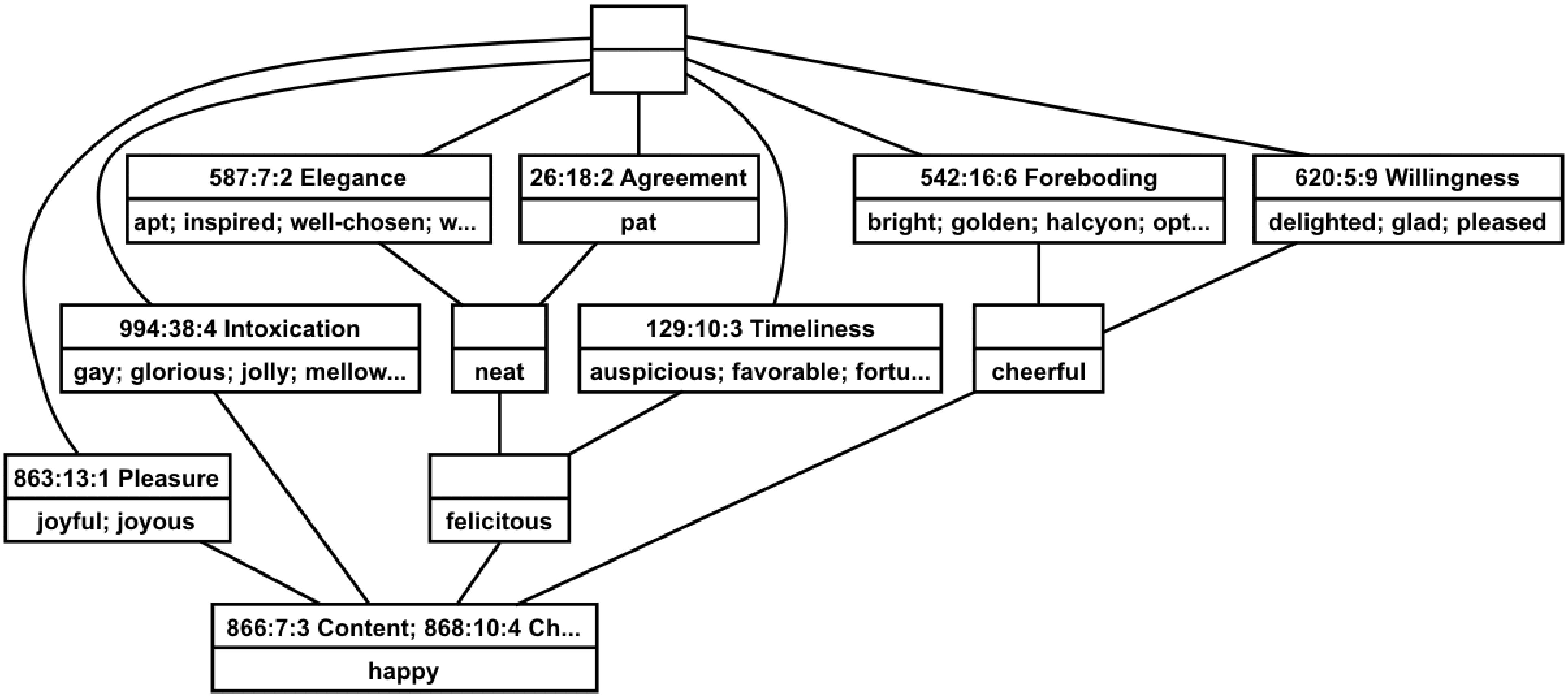
Applications:

- Ontology development and alignment; classification methods; machine learning; defining concepts used in other logics.
- FCA tools are often used to check whether ontologies specified in OWL and other notations are consistent.
- They can also be used to detect inconsistencies among two or more independently developed ontologies.

The FCA Homepage: http://www.upriss.org.uk/fca/fca.html

For deriving lattices from lexical resources: http://www.upriss.org.uk/papers/jucs04.pdf

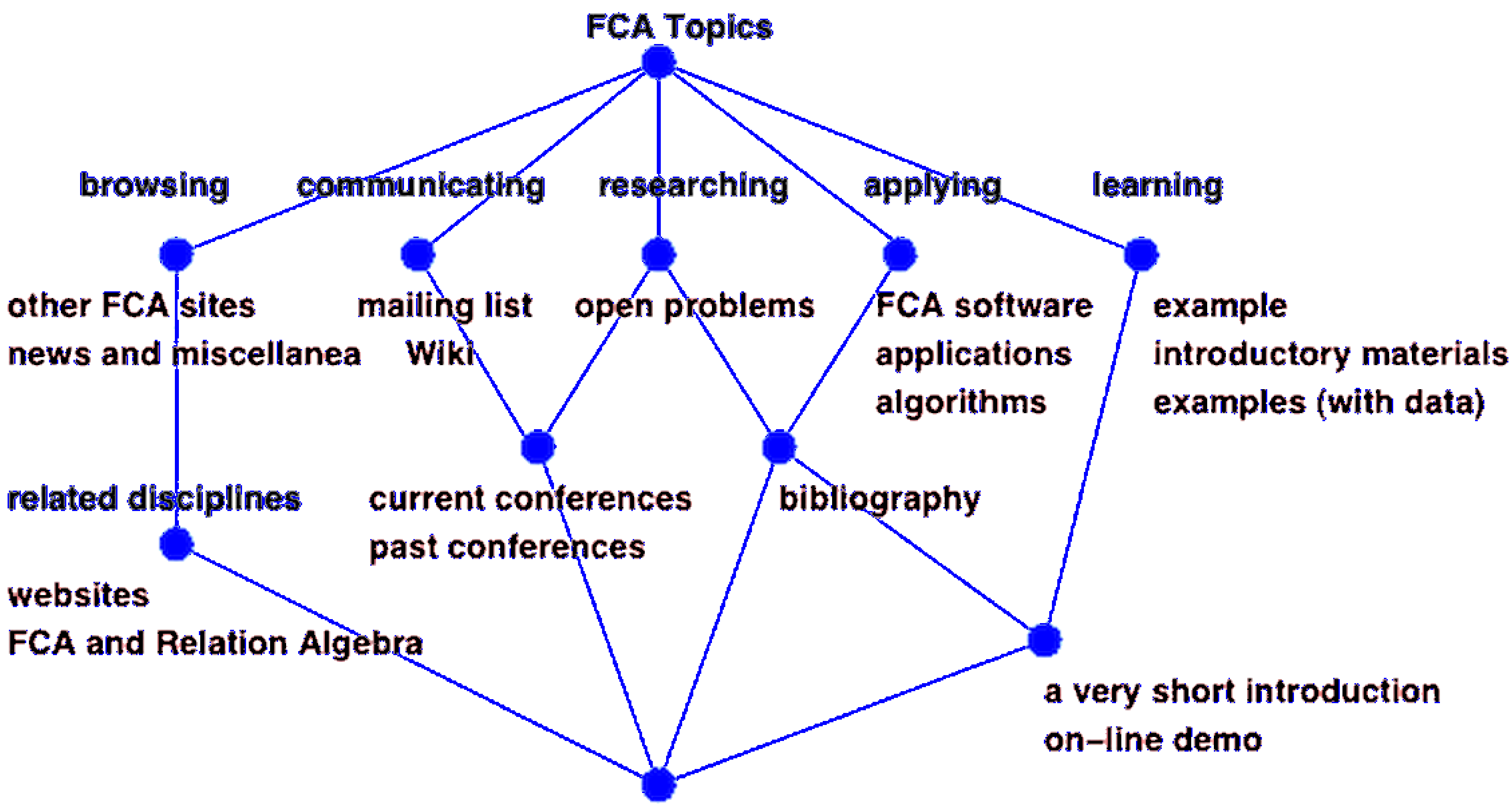
Generating Lattices Automatically



FCA tools used the data in Roget's Thesaurus to generate this lattice for the word 'happy' and its hypernyms (supertypes).

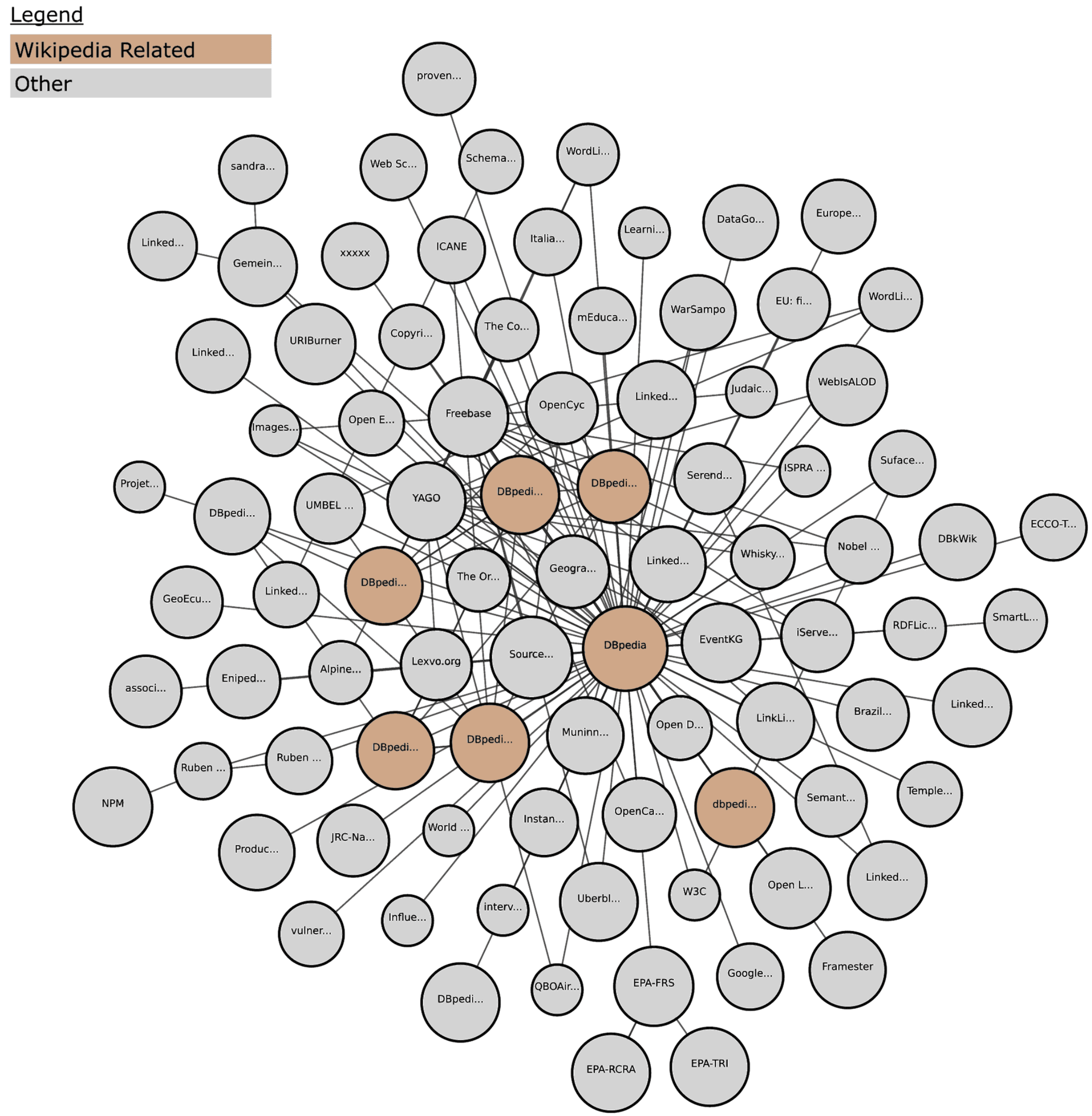
To generate this or similar lattices, enter 'happy' or any other word at the web site http://www.ketlab.org.uk/roget.html

Classifying Resources by Purpose



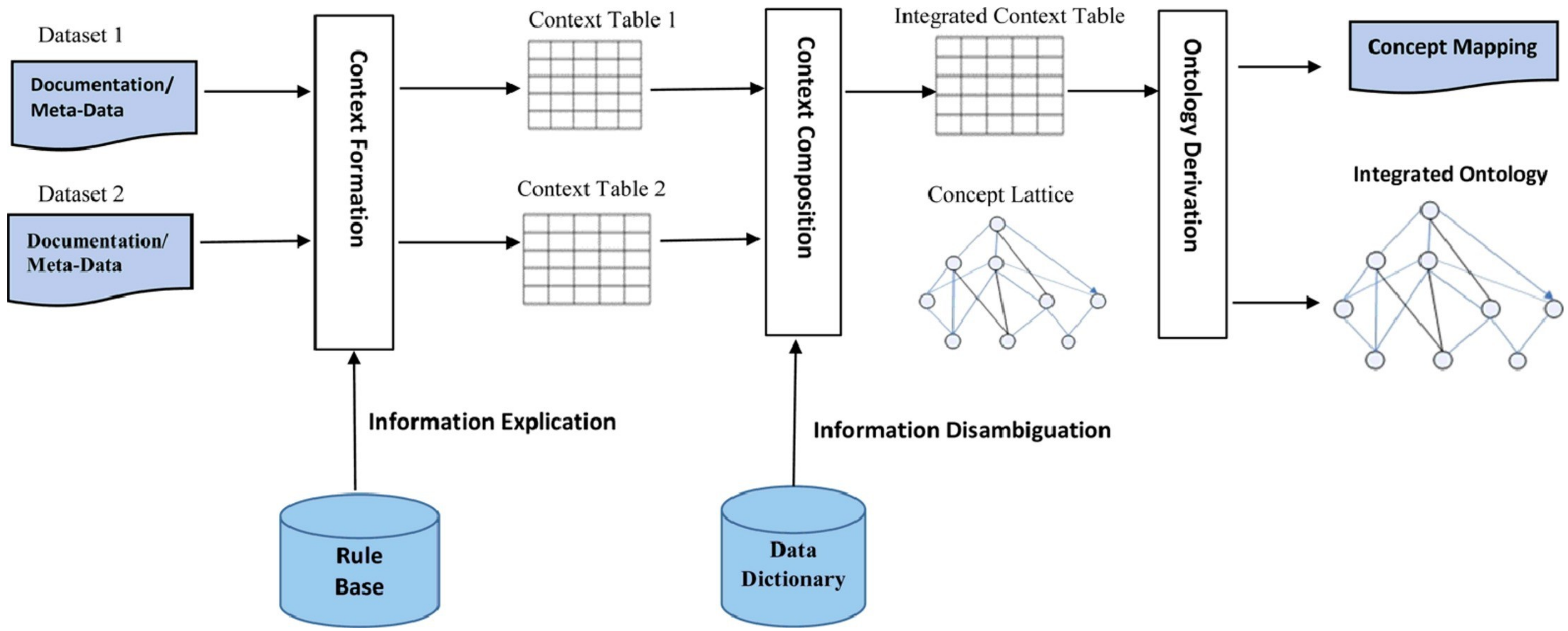
FCA tools may use a variety of criteria for classification.

- For ontology, the usual criterion is type/subtype.
- But a person who asks a question has some purpose in mind.
- The lattice above classifies resources by purpose, not type.



Challenge: Relate diverse ontologies across domains.

Using FCA to Merge Ontologies



Semi-automated method for integrating diverse ontologies.*

- **Independently developed systems are usually incompatible.**
- **FCA tools can detect similarities and conflicts in definitions.**
- **With some human assistance, the tools can derive a merged ontology that can support data sharing among the systems.**

*** Gaihua Fu, FCA based ontology development for data integration, https://www.sciencedirect.com/science/article/pii/S030645731630019X**

Federated Ontologies

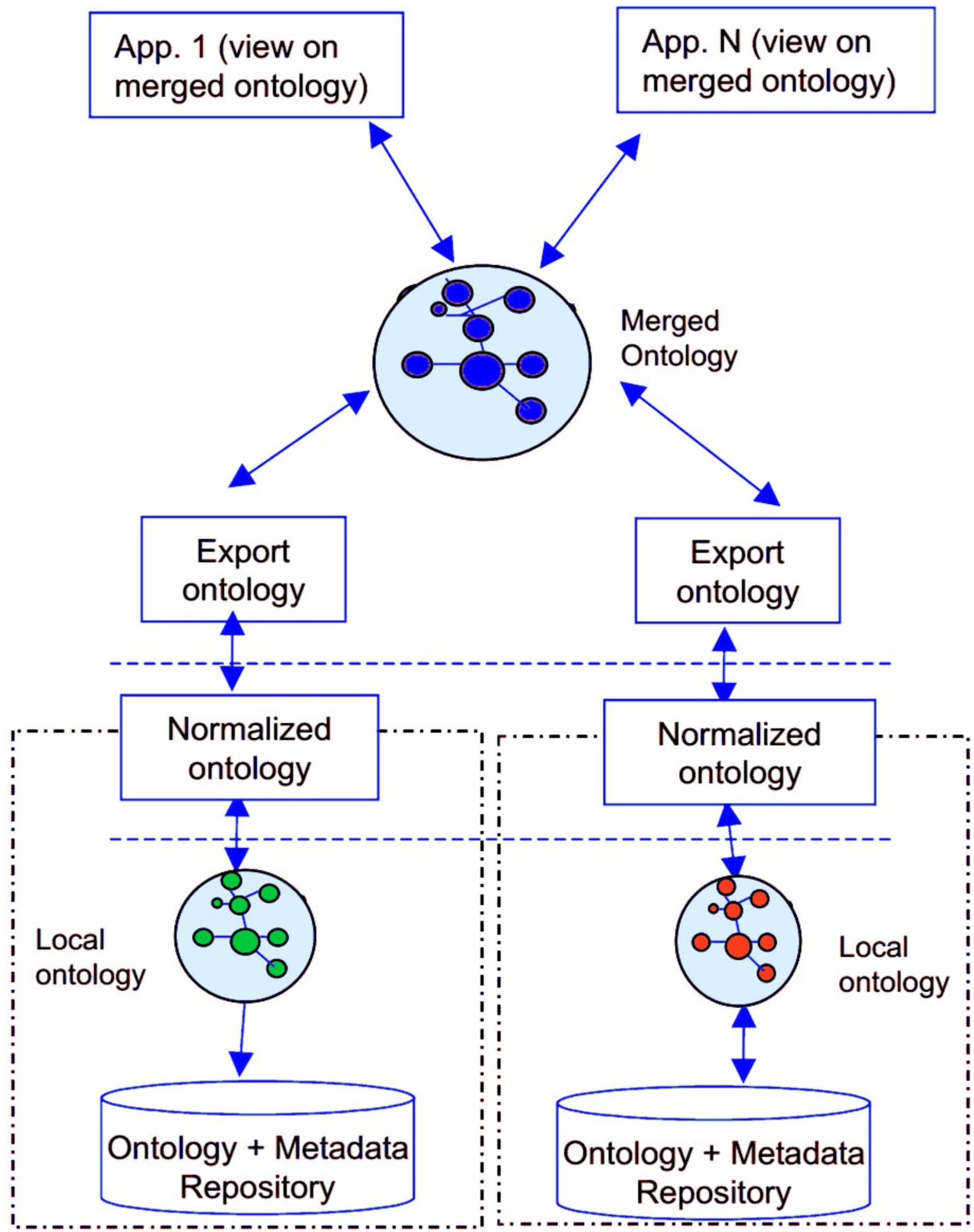
A merger of multiple ontologies may become large and unwieldy.

A federation of ontologies for a large domain would require an underspecified top level. The subdomains would add detail as needed for various applications.

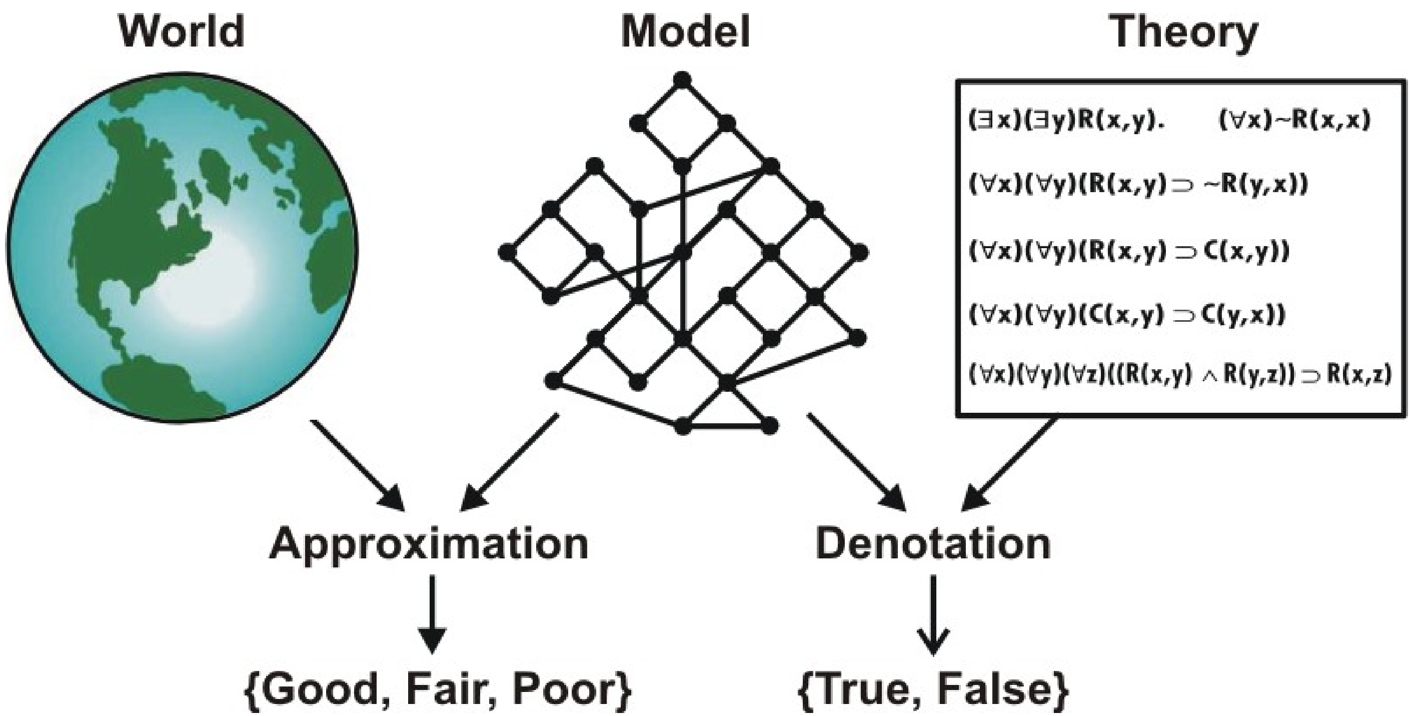
A physician, a pharmacist, a nurse, and a patient, for example, may talk about the same case, but they would discuss different details from different points of view.

The diagram on the right illustrates an application of FCA methods.

See G. Stumme & A. Maedche, Ontology merging for federated ontologies on the Semantic Web.



Relating Models to the World



Engineers: "All models are wrong, but some are useful."

- **Discrete symbolic models can be clear, sharp, and precise.**
- **But the world is continuous, disordered, and fuzzy.**

Natural languages are flexible. They can adapt to anything.

- **They can be as vague or precise as the situation requires.**
- **SW tools should be flexible: Detailed levels must be precise, but the ontology must accommodate anything imaginable.**

Knowledge of Good and Evil

Observation by Immanuel Kant:

> ***Socrates said he was the midwife to his listeners, i.e., he made them reflect better concerning that which they already knew, and become better conscious of it. If we always knew what we know, namely, in the use of certain words and concepts that are so subtle in application, we would be astonished at the treasures contained in our knowledge...***
>
> ***Platonic or Socratic questions drag out of the other person's cognitions what lay within them, in that one brings the other to consciousness of what he actually thought.***

From his ***Vienna Logic***

C. S. Peirce: ***Logic is a sort of tree of knowledge of good and evil which costs the loss of paradise to him who tastes of its fruit.***

But good tools may help us "drag out" the treasures and the treachery hidden in the sources of our knowledge.

Related Readings

ISO/IEC standard 24707 for Common Logic, http://standards.iso.org/ittf/PubliclyAvailableStandards/c039175_ISO_IEC_24707_2007(E).zip

Majumdar, Arun K., John F. Sowa, & John Stewart (2008) Pursuing the goal of language understanding, http://www.jfsowa.com/pubs/pursuing.pdf

Majumdar, Arun K., & John F. Sowa (2009) Two paradigms are better than one and multiple paradigms are even better, http://www.jfsowa.com/pubs/paradigm.pdf

Majumdar, Arun K., & John F. Sowa (2018) Relating language, logic, and imagery, http://jfsowa.com/pubs/relating.pdf

Petarinen, Ahti-Veikko (2011) Moving pictures of thought II, *Semiotica* 186:1-4, 315–331, http://www.helsinki.fi/~pietarin/publications/Semiotica-Diagrams-Pietarinen.pdf

Putnam, Hilary (1982) Peirce the logician, http://jfsowa.com/peirce/putnam.htm

Sowa, John F. (2006) Peirce's contributions to the 21st Century, http://www.jfsowa.com/pubs/csp21st.pdf

Sowa, John F. (2006) Worlds, models, and descriptions, http://www.jfsowa.com/pubs/worlds.pdf

Sowa, John F. (2010) Role of logic and ontology in language and reasoning, http://www.jfsowa.com/pubs/rolelog.pdf

Sowa, John F. (2011) Peirce's tutorial on existential graphs, http://www.jfsowa.com/pubs/egtut.pdf

Sowa, John F. (2013) What is the source of fuzziness? http://jfsowa.com/pubs/fuzzy.pdf

Sowa, John F. (2013) From existential graphs to conceptual graphs, http://www.jfsowa.com/pubs/eg2cg.pdf

Sowa, John F. (2018) Reasoning with diagrams and images, http://www.collegepublications.co.uk/downloads/ifcolog00025.pdf