Semantics of Natural Languages

John F. Sowa

22 March 2013

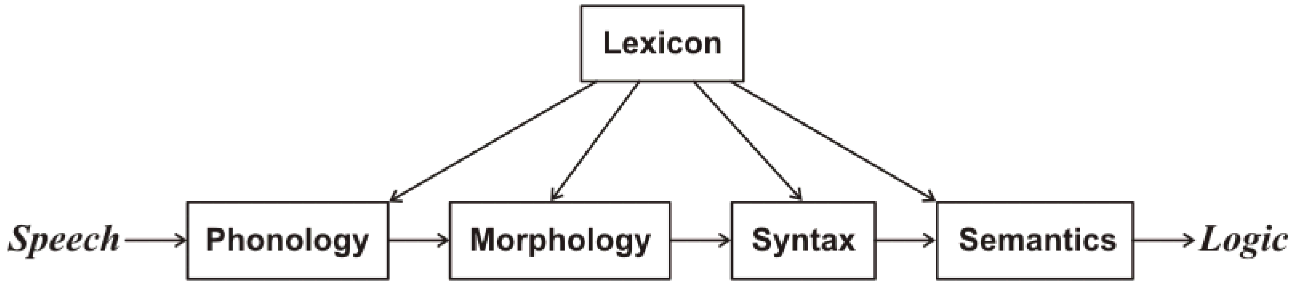
1. Controversies About Language

Ten linguists and philosophers who span the range of ideas:

- **Noam Chomsky: Generative syntax is the essence of language.**
- **Roman Jakobson: Syntax without semantics is meaningless.**
- **Michael Halliday: Language is social semiotic.**
- **Jerry Fodor: Speech is generated from a language of thought.**
- **Lotfi Zadeh: The logic of natural language is fuzzy.**
- **Richard Montague: Any natural logic must be a formal logic.**
- **Ludwig Wittgenstein: Games (*Sprachspiele*) are the foundation.**
- **Yorick Wilks: Wittgenstein was right, but more detail is needed.**
- **Roger Schank: Background knowledge is essential.**
- **Fred Jelinek: Statistics is key to all language processing.**

Can their insights tell us what kind of reasoning is natural?
What combinations would be the most natural? Computable?

Classical Natural Language Processing

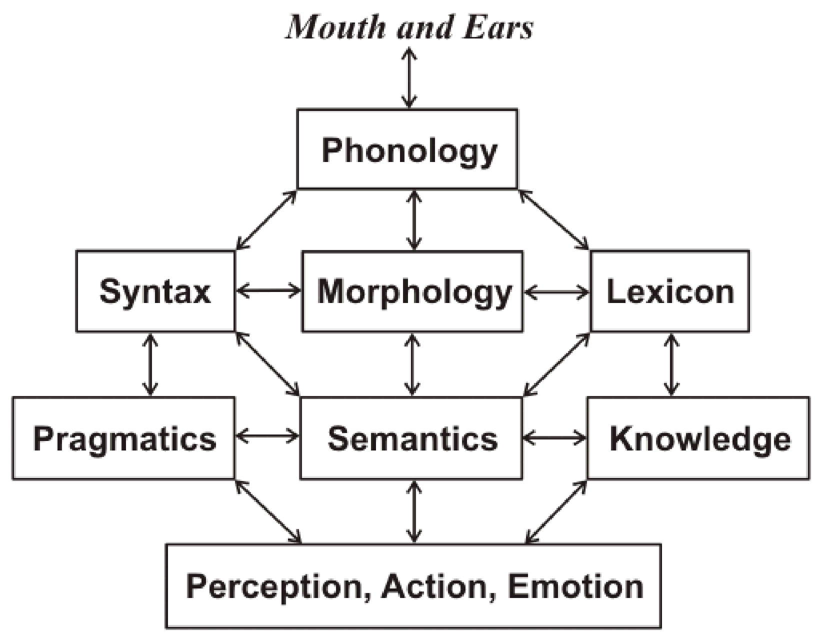


Assumptions about language that led to this approach:

- **Frege's principle of compositionality.**
- **Chomsky's distinction between competence and performance.**
- **Tarski-Kripke-Montague's model-theoretic semantics.**
- **A formal ontology that defines all word senses.**
- **A linear flow of information from phonology to semantics.**

This method can "understand" a restricted or controlled natural language, but not the usual unrestricted language.

Psychology and Neuroscience



Everything at every stage can and does interact with everything else.

Informal methods that use background knowledge at an early stage have been promising (Schank 1982, Wilks & Fass 1992).

But they require better tools for acquiring and using knowledge.

Four Views by Logicians

1. **Montague declared that NL semantics is formal:**

   "There is in my opinion no important theoretical difference between natural languages and the artificial languages of logicians; indeed, I consider it possible to comprehend the syntax and semantics of both kinds of languages within a single natural and mathematically precise theory."

2. **Frege, Russell, and Carnap believed in a sharp dichotomy:**

   "a task of philosophy [is] to break the power of words over the human mind, by uncovering illusions that through the use of language often almost unavoidably arise concerning the relations of concepts, by freeing thought from the taint of ordinary linguistic means of expression." Gottlob Frege

3. **Peirce emphasized continuity:**

   "Symbols grow. They come into being by development out of other signs, particularly from icons, or from mixed signs partaking of the nature of icons and symbols."

4. **Wittgenstein moved from view #1 to a variant of #3:**

   "new types of language, new language-games, as we may say, come into existence, and others become obsolete and get forgotten."

Precision and Ambiguity

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."

Many linguists:

Lexical ambiguity can be avoided within a semantically restricted sublanguage, such as weather reports or airplane reservations.

Ludwig Wittgenstein:

Outside a language game, a word is like "a wheel turning idly."

Words acquire a precise, unambiguous meaning only in a narrow domain, such as a sublanguage or language game.

2. Psycholinguistics

Human language is based on the way people think about everything they see, hear, feel, and do.

And thinking is intimately integrated with perception and action.

The semantics and pragmatics of a language are

- **Distributed in the brains of every speaker of the language,**
- **Situated in time and space,**
- **Dynamically generated and interpreted in terms of a constantly developing and changing context,**
- **Embodied and supported by the sensory and motor organs.**

These points summarize current views by psycholinguists.

Philosophers and logicians have debated other issues:

- **NL as a formal logic; a sharp dichotomy between NL and logic; a continuum between NL and logic.**

Discrete and Continuous

The world is a continuum, but words are discrete.

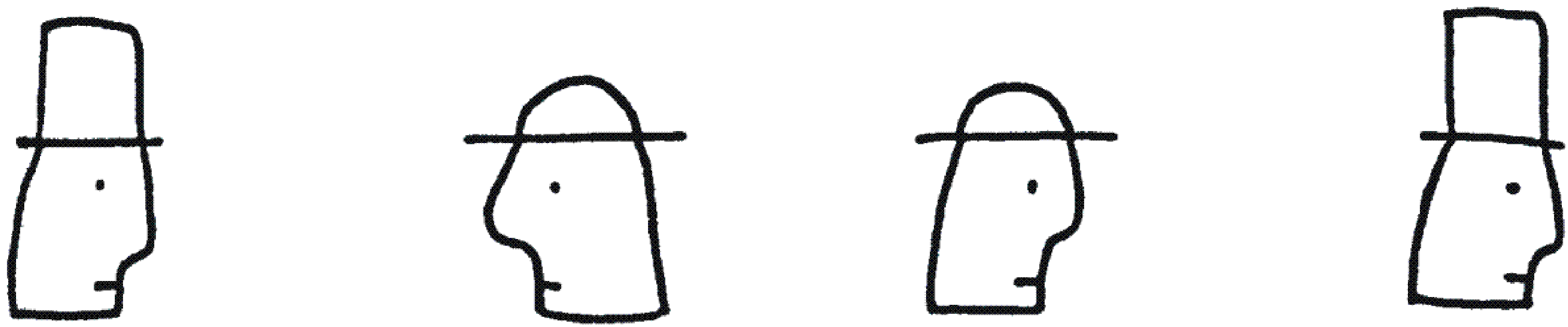
- **How can discrete words describe a continuous world?**
- **Short answer: Allow a continuous variation in meaning.**

But why are words discrete?

- **You can't build large stable structures with unstable stuff.**
- **You can build a mud hut, but not a mud skyscraper.**
- **Complex designs are built with discrete building blocks: DNA, the cells of living things, and large engineering projects.**
- **Most animals communicate with a small number of graded (continuously variable) sounds.**
- **The discrete words and phonemes of human language enable an infinite variety of documents of any size: Look at the WWW.**

Methods of perception, action, learning, reasoning, and communication must relate the discrete and the continuous.

Relating Language to Perception



le curé est devant le ministre
the priest is in front of the minister

le curé est derrière le ministre
the priest is behind the minister

Vandeloise drew diagrams to explain spatial terms in French.

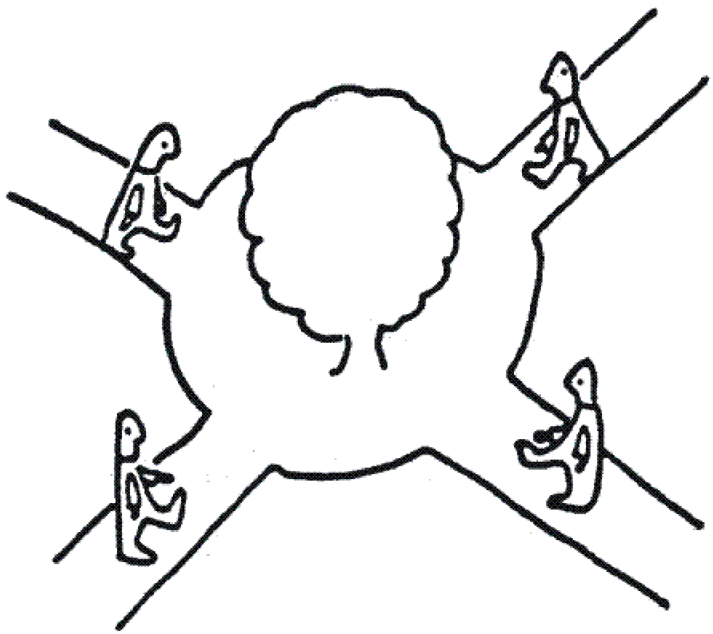
For most words, dictionaries list many meanings.

But the number of possible meanings is open ended.

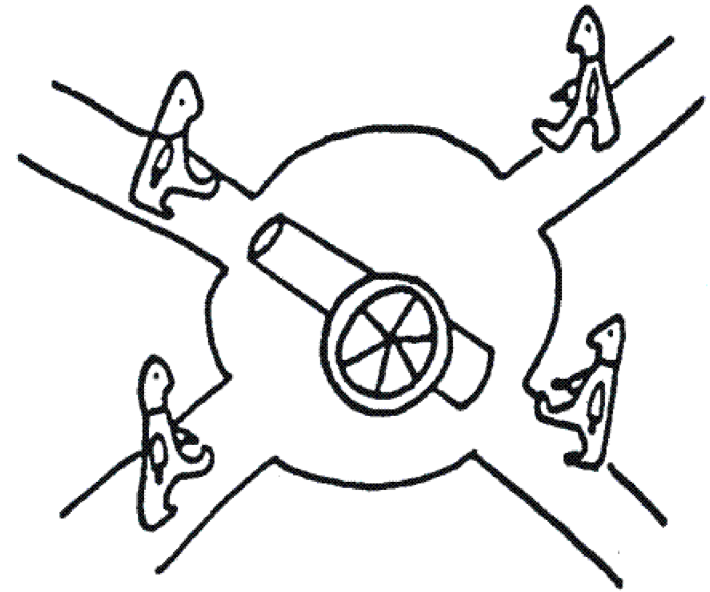
Even the claim of a core literal meaning is doubtful.

Diagrams adapted from Vandeloise (1986)

Issues about Orientation



les présidents sont devant l'arbre
the presidents are in front of the tree

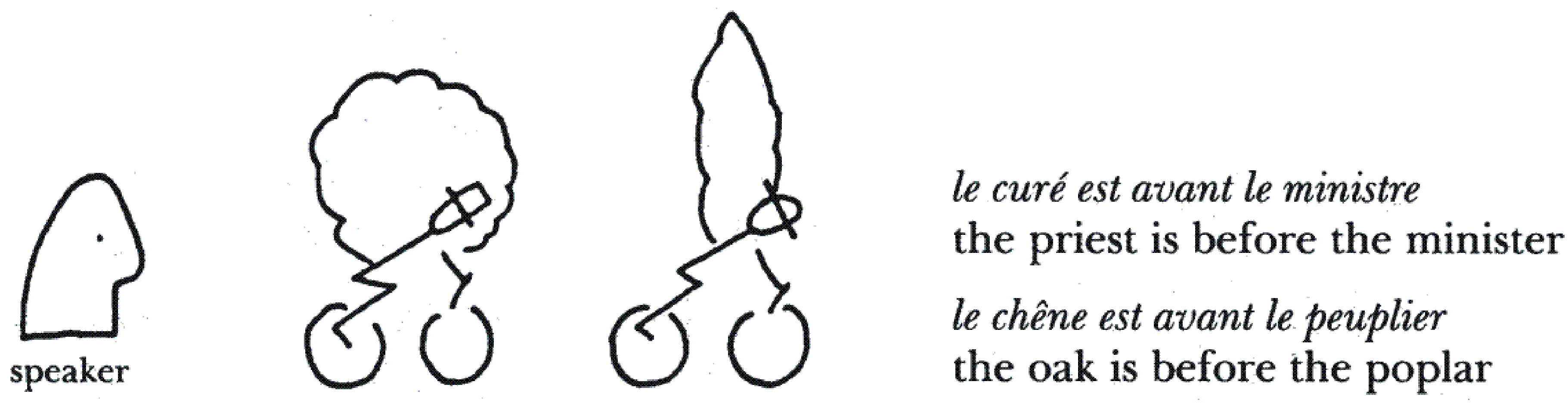


? les présidents sont devant le canon
the presidents are in front of the cannon

For a tree, any side could be could be considered the front.

But a cannon has distinct front, back, and sides.

Issues about Motion



For stationary objects, such as trees, the speaker's viewpoint determines the choice of preposition.

For moving objects, their relative position is more significant.

But objects like snails and turtles, which move very slowly, are treated like stationary objects (unless their motion is relevant).

Issues about Function

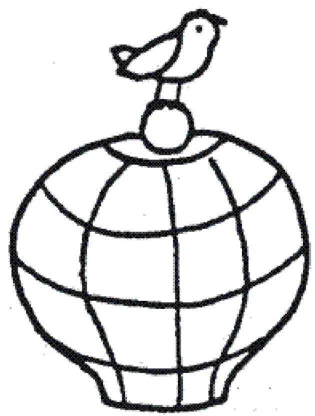
la poire est dans la coupe
the pear is in the bowl

The French preposition *dans* or the English *in* normally links something to a container.

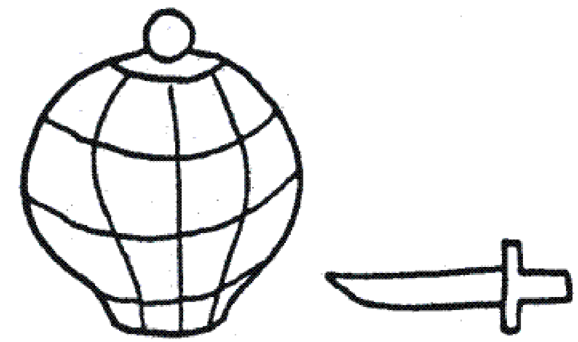
The primary function of a bowl is to serve as a container.

That function is more relevant than the question whether the bowl actually encloses the pear.

Issues about Background Knowledge



l'oiseau est à l'extérieur de la cage
the bird is outside the cage



? le couteau est à l'extérieur de la cage
the knife is outside the cage

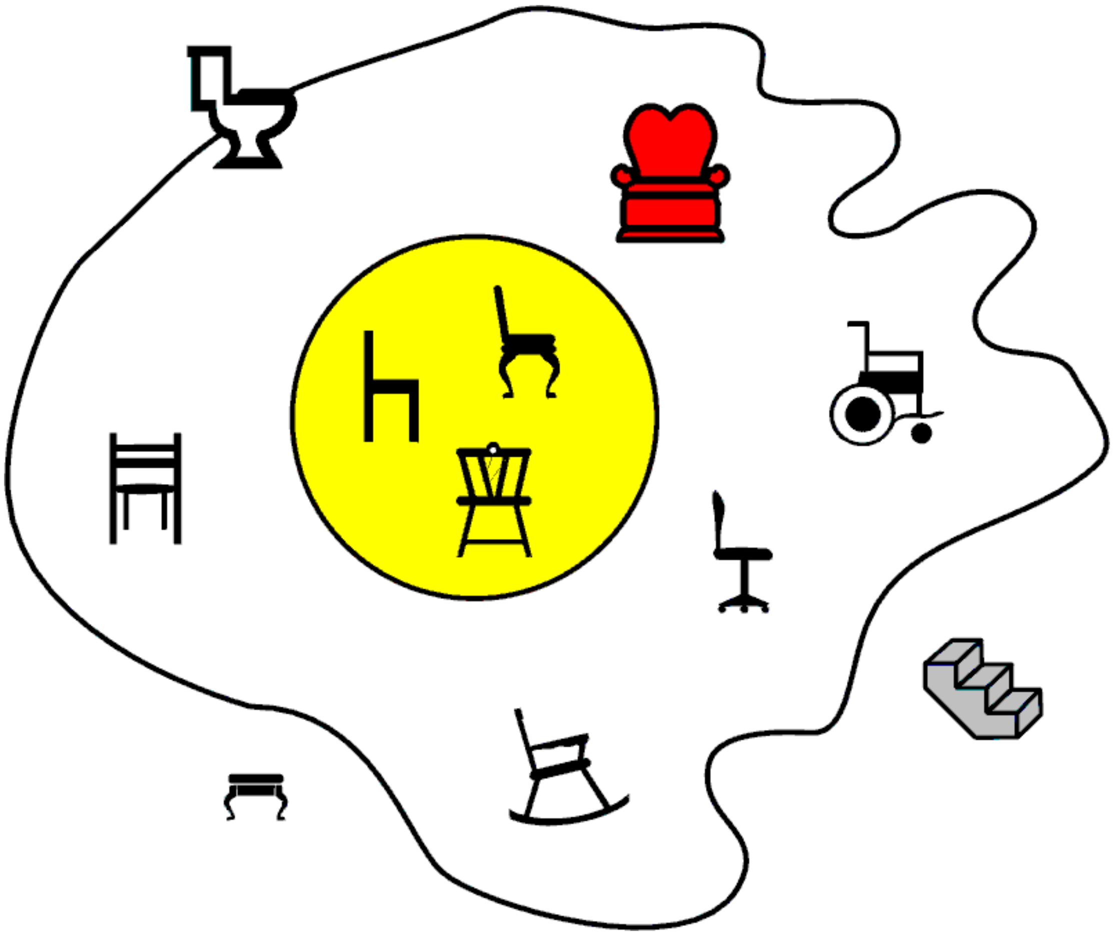
A cage is sometimes used to enclose a bird.

But a cage is an unlikely container for a knife.

Normal comment: "The knife is to the right of the cage."

To say "The knife is outside the cage" implies that there is some reason why it might have been in the cage.

What is a Chair?



The egg-yolk theory puts typical examples in the yolk and unusual examples in the egg white (Lehmann & Cohn 1994).

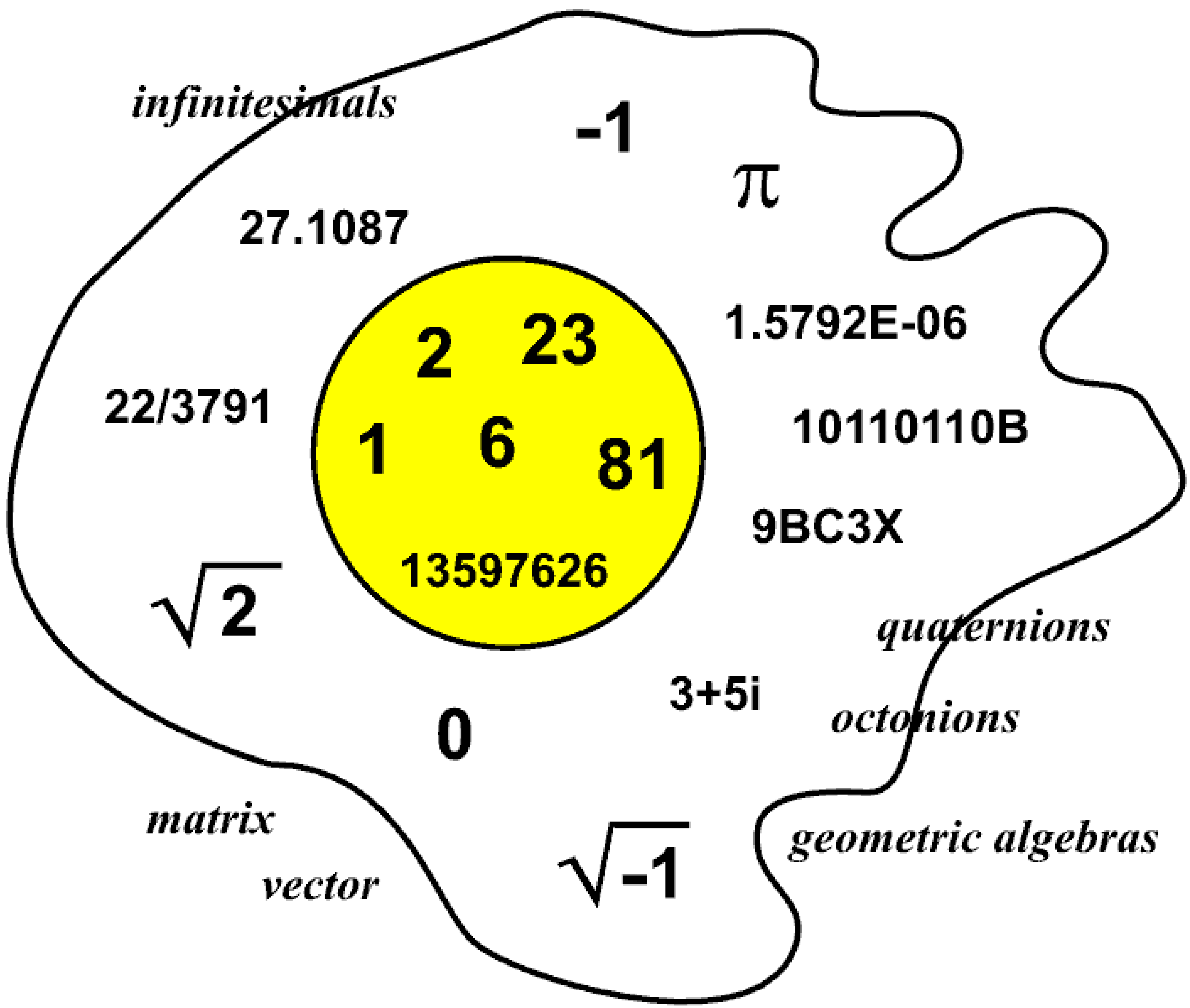
The boundaries in this example are similar to the level cuts that partition a fuzzy set (Bandler & Kohout 1988).

Is it a Chair? Art? Humor? Fantasy?

In a museum, it's funny. But what if you saw it at night, alone, in an old castle. Meaning always depends on context.

Claw and Ball Chair by Jake Cress. At the Smithsonian Renwick Gallery.

What is a Number?



Concepts in science and mathematics grow and change.

Microsenses

The linguist Allen Cruse coined the term *microsense* for a specialized sense of a word in a particular application.

Examples of microsenses:

- **Spatial terms in different situations and points of view.**
- **The many kinds of chairs or numbers in the egg whites.**
- **The kinds of balls in various ball games: baseball, basket ball, billiard ball, bowling ball, football, golf ball, softball, tennis ball.**
- **Computer science requires precise definitions, but the meanings change whenever programs are revised or extended.**
- **Consider the term *file system* in Unix, Apple OS X, Microsoft Windows, and IBM mainframes.**

Microsenses develop through usage in different situations.

The number and kinds of new uses and innovations grow independently of any attempt to limit the meanings of words.

“I don’t believe in word senses.”

The title is a quotation by the lexicographer Sue Atkins, who devoted her career to writing and analyzing word definitions.

In an article with that title,* Adam Kilgarriff observed that

- **“A task-independent set of word senses for a language is not a coherent concept.”**
- **The basic units of meaning are not the word senses, but the actual “occurrences of a word in context.”**
- **“There is no reason to expect the same set of word senses to be relevant for different tasks.”**
- **“The set of senses defined by a dictionary may or may not match the set that is relevant for an NLP application.”**
- **Professional lexicographers are well aware of these issues.**
- **The senses they select for a dictionary entry are based on editorial policy and assumptions about the readers’ expectations.**

*** See http://www.kilgarriff.co.uk/Publications/1997-K-CHum-believe.pdf**

Using Background Knowledge

People resolve ambiguities and choose the correct microsenses by retrieving background knowledge about the options.

Choosing the microsense: *My dog bit the visitor's ear*.

- **From knowledge about the size of dogs, one would assume it was more likely to be a doberman than a dachshund.**
- **But if one knew the visitor was in the habit of bending over to pet a dog, it might even be a chihuahua.**

Resolving an ambiguous parse: *The chicken is ready to eat*.

- **From knowledge about typical food, one would assume the chicken had been cooked and prepared as a meal.**
- **If the word** *chicken* **were replaced with** *dog*, **one might assume the dog was begging for food.**
- **But people in different cultures may make different assumptions.**

The many microsenses and the dependence on background knowledge require highly flexible methods of reasoning.

Evolution of Grammar

Grammatical forms – both inflections and phrase patterns – are as ambiguous as words and have just as many microsenses.

Study by Bybee, Perkins, and Pagliuca (1994):

- **Analysis of 76 languages from every major language group in the world.**
- **Grammar evolves from frequently occurring semantic patterns that become stylized and generalized to express a useful distinction.**
- **Over time, sound changes simplify the forms, they acquire too many microsenses, and they become too ambiguous to be useful.**
- **New patterns evolve to make more precise distinctions.**

Example: ***be going to*** **as a marker of future tense.**

- **Originally expressing motion toward a goal:** ***She's going to Boston*****.**
- **Adding a verb to qualify the motion:** ***She's going to drive to Boston*****.**
- **Generalizing to an abstract goal:** ***She's going to write a book*****.**
- **Simplifying and overgeneralizing:** ***She's gonna have a baby*****.**

The Ultimate Understanding Engine

Sentences uttered by a child named Laura before the age of 3. *

Here's a seat. It must be mine if it's a little one.

I went to the aquarium and saw the fish.

I want this doll because she's big.

When I was a little girl, I could go "geek geek" like that, but now I can go "This is a chair."

Laura used a larger subset of logic than Montague formalized.

No computer system today has Laura's ability to speak and understand language.

*** John Limber, The genesis of complex sentences. In T. Moore (Ed.), *Cognitive development and the acquisition of language*. New York: Academic Press, 1973.**
http://pubpages.unh.edu/~jel/JLimber/Genesis_complex_sentences.pdf

Evolution of Logic

At age 3, Laura correctly used the logic words of English, but it's unlikely that she mapped them to and from a formal logic.

How does logic develop in a child or in the species?

- **Languages have words like *and, or, not, if, some,* and *every*.**
- **But those words have as many context-dependent microsenses as any other words or grammatical forms.**
- **That multiplicity rules out a unique "natural logic" or "language of thought" with strict rules of inference.**
- **Instead, logics are like "games" that develop in the same way as games like chess and bridge or any version of mathematics.**
- **But logics have strong similarities to language.**

How could logic be supported by the same mechanisms of perception and action as language?

3. Ludwig Wittgenstein

Considered one of the greatest philosophers of the 20th century.

Wrote his first book under the influence of Frege and Russell.

That book had an enormous influence on analytic philosophy, formal ontology, and formal semantics of natural languages.

But Wittgenstein retired from philosophy to teach elementary school in an Austrian mountain village.

In 1929, Russell and others persuaded him to return to Cambridge University, where he taught philosophy.

During the 1930s, he began to rethink and criticize the foundations of his earlier book, including many ideas he had adopted from Frege and Russell.

Wittgenstein's First Book

From the *Tractatus Logico-Philosophicus,*

1 The world is everything that is the case.

1.1 The world is the totality of facts, not of things.

3.25 There is one and only one complete analysis of the proposition.

4.001 The totality of propositions is the language.

4.116 Everything that can be said can be said clearly.

5 Propositions are truth-functions of elementary propositions.

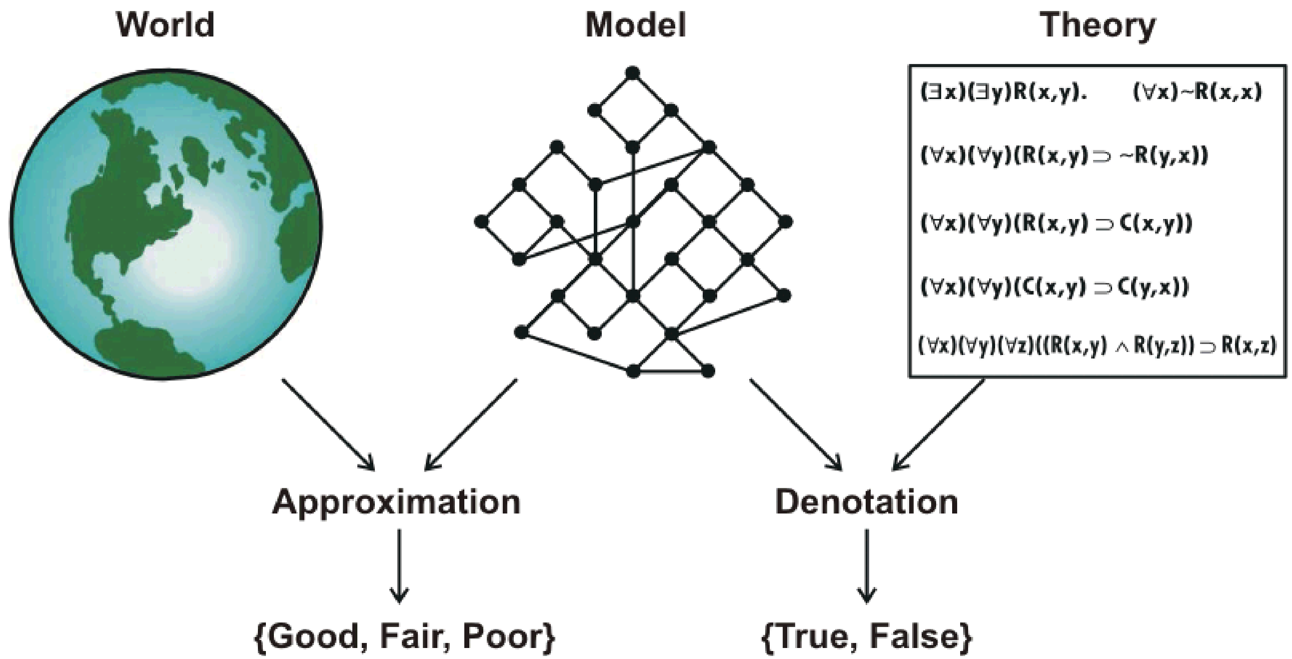
6.13 Logic is not a theory but a reflexion of the world.

7 Whereof one cannot speak, thereof one must be silent.

This book set the agenda for formal semantics in the 20th century.

If it were adequate for language understanding and reasoning, then the HAL 9000 would be ruling the world today.

Model-Theoretic Semantics



In the *Tractatus*, Wittgenstein assumed that the world is the model.

If there is exactly one world, there is exactly one model, there is exactly one ontology, and any approximation is false or meaningless.

Engineers are cynical, but realistic:

"All models are wrong. Some are useful."

What Kind of Language is Meaningless?

According to the *Tractatus*, only factual statements about the world and Boolean combinations of them are meaningful.

> **6.421 It is clear that ethics cannot be expressed. Ethics is transcendental. (Ethics and aesthetics are one.)**
>
> **6.52 We feel that even if all possible scientific questions be answered, the problems of life have still not been touched at all. Of course there is then no question left, and just this is the answer.**
>
> **6.54 My propositions are elucidatory in this way: he who understands me finally recognizes them as senseless, when he has climbed out through them, on them, over them.**

Since the *Tractatus* consists of language about language, it does not state facts about the world. Therefore, it too is meaningless.

Wittgenstein's Transitional Period

In 1929-30, he analyzed some "minor" inconsistencies in the *Tractatus*.

That analysis led to two important innovations:

- **Satzsystem: System of sentences or propositions.**
- **Beweissystem: Proof system that defines a logic for a Satzsystem.**

This approach distinguishes the model from the world:

> **"The Satzsystem is like a ruler laid against reality. An entire system of propositions is now compared to reality, not a single proposition."**

For a given logic, each consistent Satzsystem expressed in that logic is a theory that defines an ontology.

The model is no longer identical to the world, and different Satzsysteme could be better or worse approximations for different purposes.

Uses and Limitations of Satzsysteme

In the transitional period, Wittgenstein continued his focus on logic:

- He allowed multiple logics and ontologies, and he relaxed the truth criteria to map systems to the world, not single sentences.
- He did not discuss ethics, aesthetics, or metalanguage.
- But by distinguishing models from the world, he allowed models to represent features that have no direct mapping to the world.

This approach allows a broader range of topics.

It can be useful as a semantic basis for controlled NLs.

But it is inadequate for unrestricted natural language.

Wittgenstein's Language Games

In the mid 1930s and for the rest of his life, Wittgenstein focused on his theory of language games as a more general and flexible approach than the Satzsysteme and Beweissysteme of 1929-30.

In his book *Philosophical Investigations*, he presented them as a correction to the "grave errors" (schwere Irrtümer) of his first book.

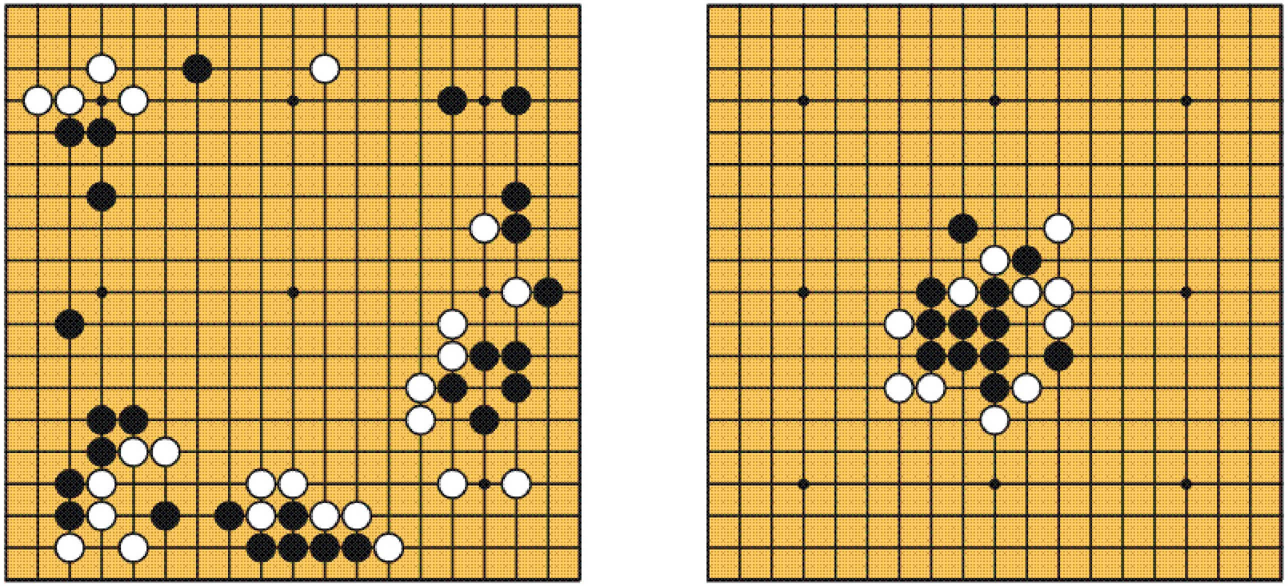
Every use of language is intimately integrated with social activity. The meaning of language is grounded in purposeful activity:

> **"The word 'language-game' is used here to emphasize the fact that the *speaking* of language is part of an activity, or of a form of life." (§23)**

As an example, he said that a word is like a piece in the game of chess. Its meaning is based on the way the word is used in a game.

Games of Go and Go-moku

Two games with the same syntax, but different goals:



Syntax defines legal moves, but not meaningful moves.

The meaning of any move is determined by its purpose.

In go, the goal is to place stones that surround territory.

In go-moku, the goal is to place five stones in a row.

Different goals lead to very different patterns of stones.

Wittgenstein's Examples

"Consider the variety of language games in the following examples, and in others:

Giving orders, and acting on them –
Describing an object by its appearance, or by its measurements –
Constructing an object from a description (a drawing) –
Reporting an event –
Speculating about the event –
Forming and testing a hypothesis –
Presenting the results of an experiment in tables and diagrams –
Making up a story; and reading one –
Acting in a play –
Singing rounds –
Guessing riddles –
Making a joke; telling one –
Solving a problem in applied arithmetic –
Translating from one language into another –
Requesting, thanking, cursing, greeting, praying." (§23)

Implementing Language Games

Wittgenstein's language games are still controversial, partly because they cross many academic boundaries:

syntax, semantics, pragmatics, logic, ontology, speech acts, themes, narratives, scenarios, sublanguage, genre...

Many logicians consider them "a step in the wrong direction" away from the clarity and precision of his first book.

Some computational linguists found his writings inspirational, but they found it difficult to implement that inspiration.

The Satzsysteme are easier to formalize, and they could be considered special cases of language games.

But how could they be formalized and implemented?

A Neo-Wittgenstenian Model of Language

Developed by Margaret Masterman —

One of six students in Wittgenstein's course of 1933-34 whose notes were compiled as *The Blue Book*.

Founded the Cambridge Language Research Unit (CLRU) as a discussion group that became a major center for NLP.

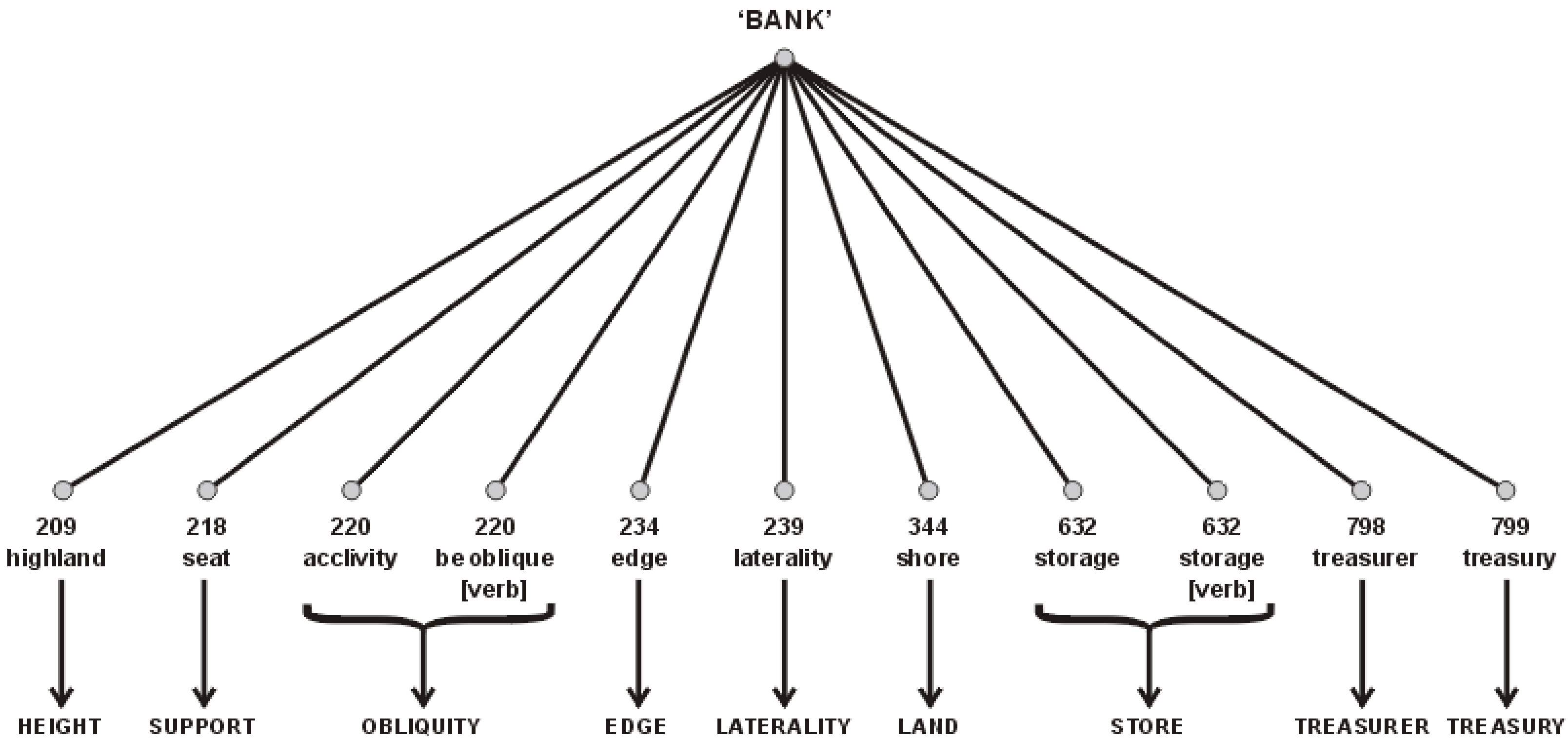
Emphasized semantics, not syntax:

> **"I want to pick up the relevant basic-situation-referring habits of a language in preference to its grammar."**

Invented a context-dependent method of analysis:

1. **Thesaurus with words grouped by areas of use.**
2. **Word "fans" radiating from each word type to each area of the thesaurus in which it occurs.**
3. **Dynamically generated combinations of fans for word tokens.**

A Word Fan for "Bank"



Numbers and labels represent areas in Roget's Thesaurus.

Method of Disambiguation

Example: "up the steep bank" and "in the savings bank".

All the words except "the" have word fans.

Masterman's method for selecting an appropriate word sense:

- **Combinations of fans "pare down" the ambiguities "by retaining only the spokes that retain ideas which occur in each."**
- **For this example,**
  - **— OBLIQUITY 220 is common to 'STEEP' and 'BANK'.**
  - **— STORE 632 and TREASURY 799 are common to 'SAVINGS' and 'BANK'.**

This method uses background knowledge to resolve ambiguities.

It's a step in the right direction, but it's too static to support the dynamic interactions of Wittgenstein's games.

4. Dynamics of Language and Reasoning

Natural languages adapt to the ever-changing phenomena of the world, the progress in science, and the social interactions of life.

No computer system is as flexible as a human being in learning and responding to the dynamic aspects of language.

Three strategies for natural language processing (NLP):

1. **Neat: Define formal grammars with model-theoretic semantics that treat NL as a version of logic. Wittgenstein pioneered this strategy in his first book and became the sharpest critic of its limitations.**
2. **Scruffy: Use heuristics to implement practical applications. Schank was the strongest proponent of this approach in the 1970s and ’80s.**
3. **Mixed: Develop a framework that can use a mixture of neat and scruffy methods for specific applications.**

NLP requires a dynamic foundation that can efficiently relate and integrate a wide range of neat, scruffy, and mixed methods.

Dynamic Semantics

To be computable, mental models, language games, and theories about them must be formally defined.

A Tarski style of model theory is too static to represent the dynamics of human thought, language, and reasoning.

The flexibility of language games and mental models requires a framework that can grow and change dynamically.

Issues in dynamic semantics:

- Relating sentences in a text or discourse to one another, to the immediate context, and to background knowledge.
- Constructing a computable model of the semantic content that can evolve during a narrative or debate.
- Dynamic methods for transforming the models to reflect changes in the world and in language, thoughts, and plans about the world.
- Evaluating the truth or relevance of the models in terms of perception of a current situation or memories of past situations.

Versions of Dynamic Semantics

Linguists and logicians developed theories about the dynamic aspects of language in discourse and narratives:

- **The first was Peirce, who called the operations on existential graphs "a moving picture of the action of the mind in thought" (Pietarinen 2005).**
- **Hintikka (1973) defined a *surface model* of a sentence S as "a mental anticipation of what can happen in one's step-by-step investigation of a world in which S is true."**
- **Kamp (1981) developed *Discourse Representation Theory* as a dynamic method of modifying a logical form to accommodate new information.**
- **Groenendijk and Stokhof (1991) developed *dynamic predicate logic* with theories and models that can be updated during discourse.**
- **Van Eijck and Visser (2010) surveyed a wide range of variants. ***

But the dynamic methods must be integrated with all forms of perception, thinking, learning, reasoning, talking, and acting.

*** See http://plato.stanford.edu/entries/dynamic-semantics/**

Discourse Representation Theory

DRT is a widely used version of dynamic semantics. *

- **Based on a logical form called Discourse Representation Structure (DRS).**
- **Has rules for building DRSes from a stream of sentences in discourse.**
- **Covers a wide variety of linguistic and logical issues including anaphora, generalized quantifiers, plurals, tense, and aspect.**
- **Defines a model-theoretic semantics for the DRS logic.**

Relating DRT to a broader range of cognitive science:

- **The structure of DRS is isomorphic to Peirce's existential graphs (EGs).**
- **Peirce and Kamp independently invented their notations while they were searching for a logical form that had a direct mapping to language.**
- **Kamp's rules for mapping language to DRS can also be adapted to EGs.**
- **The psychologist Johnson-Laird noted that Peirce's EGs and rules of inference are a good candidate for a neural theory of reasoning. ****

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

**** Philip N. Johnson-Laird (2002) Peirce, logic diagrams, and the elementary processes of reasoning, *Thinking and Reasoning* 8:2, 69-95. http://mentalmodels.princeton.edu/papers/2002peirce.pdf**

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

The rules for scope of quantifiers are different.

- **Example:** ***If a farmer owns a donkey, then he beats it.***
- **In English, quantifiers in the if-clause govern the then-clause.**
- **But in predicate calculus, the quantifiers must be moved to the front.**
- **Formula: $(\forall x)(\forall y)((\text{farmer}(x) \wedge \text{donkey}(y) \wedge \text{owns}(x,y)) \supset \text{beats}(x,y))$.**

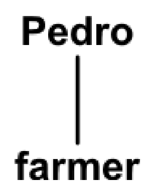
In narratives, the default operator between NL sentences is usually equivalent to '*and then*'.

Linking Operations for EG and DRS

Formulas in Peirce-Peano notation don't naturally link together.

- **Two sentences in a paragraph:** ***Pedro is a farmer. He owns a donkey.***
- **P-P formulas:** $\exists x(\mathrm{Pedro}(x) \land \mathrm{farmer}(x))$. $\exists y \exists z(\mathrm{owns}(y,z) \land \mathrm{donkey}(z))$.
- **There is no operator that can link x in the first formula to y in the second.**

Sentences translated to an EG (left) and a DRS (right):

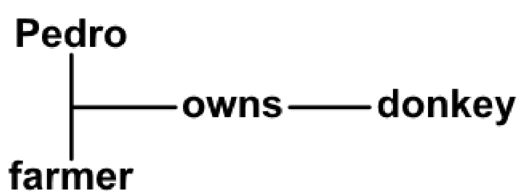


x
Pedro(x) farmer(x)

Lines in EG or variables in DRS imply existence; conjunction is implicit.

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

By connecting lines in the EGs or merging boxes in the DRSes,



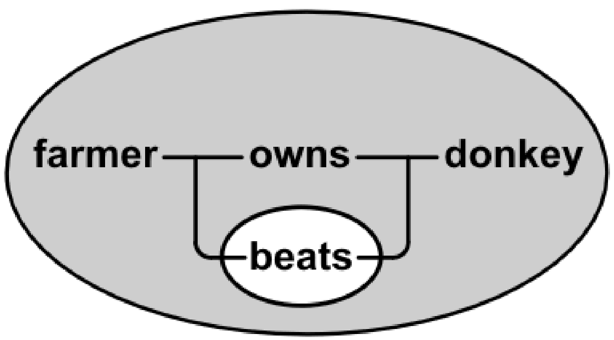
x y z x=y
Pedro(x) farmer(x)
owns(y,z) donkey(z)

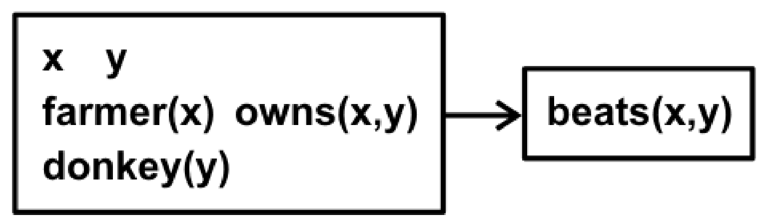
Scope of Quantifiers

For most logics, the scope of quantifiers differs from NL usage.

- **Example:** ***If a farmer owns a donkey, then he beats it.***
- **In English, quantifiers in the if-clause govern the then-clause.**
- **But in the formula, ∃ must be moved to the front and changed to ∀.**
- **Formula:** $\forall x \forall y((\text{farmer}(x) \wedge \text{donkey}(y) \wedge \text{owns}(x,y)) \supset \text{beats}(x,y))$.

The rules for EG and DRS are closer to English and other NLs:





Each oval in the EG represents negation. Any quantifier in the outer negation (shaded area) includes the inner negation within its scope.

By convention, the DRS if-box includes the then-box within its scope.

Peirce’s Rules of Inference

Three pairs of rules, which insert or erase a graph or subgraph:

- **Insert/Erase: Insert any graph in a positive (unshaded) area; erase any graph in a negative (shaded) area.**
- **Iterate/Deiterate: Iterate (copy) any graph in the same area or any nested area; deiterate (erase) any graph that could have been iterated.**
- **Double negation: Insert or erase a double negation (pair of ovals with nothing between them) around any graph in any area.**

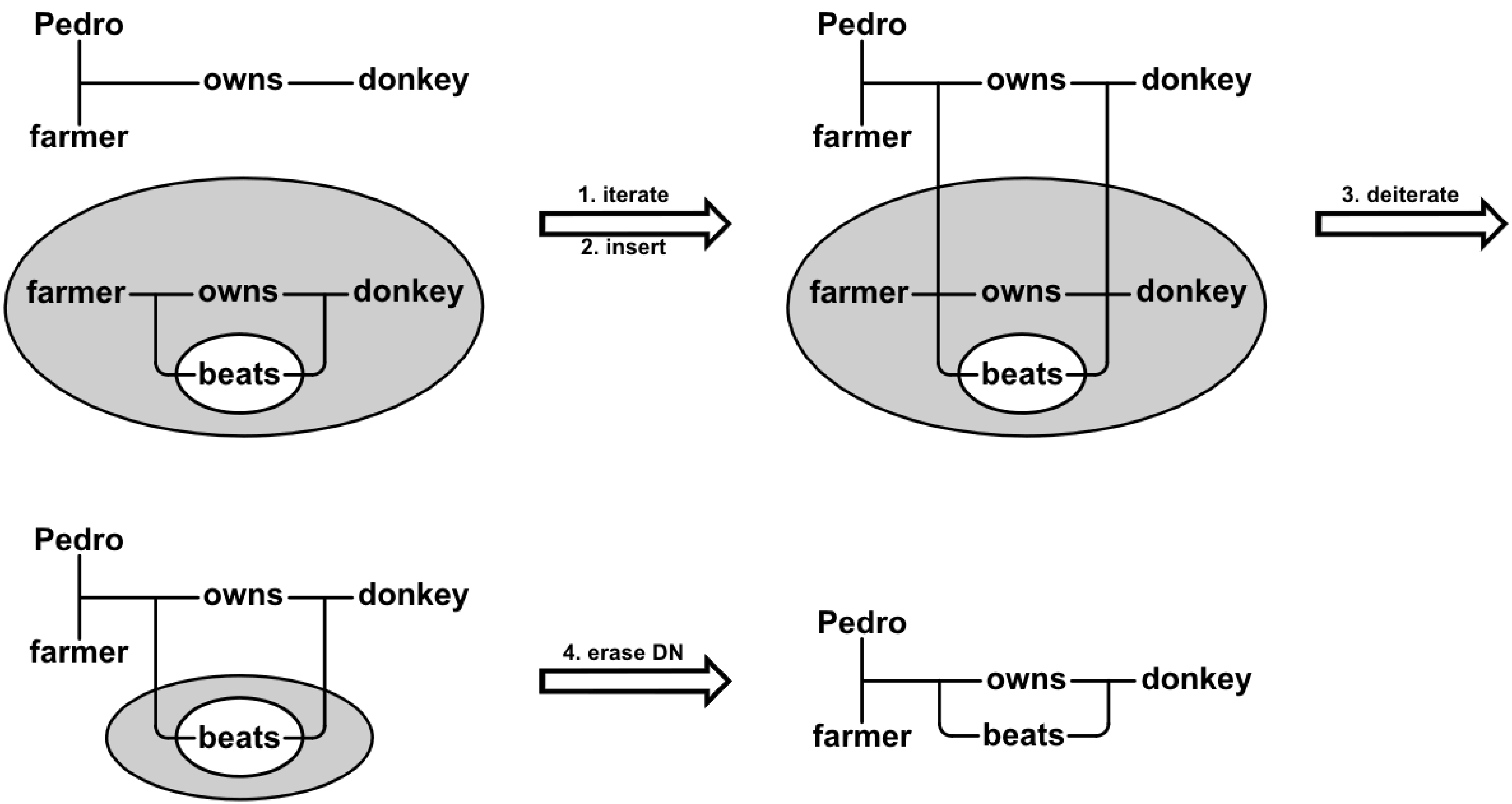
A four-step proof in the next slide:

1. **Iterate (extend) a pair of lines into a nested negative area.**
2. **Insert connections from those lines to lines in that negative area.**
3. **Deiterate (erase) a graph that is identical to one in an outer area.**
4. **Erase a double negation.**

Note that every step is a simple operation that could be performed by making or inhibiting connections between neurons. (See Lamb 2011.)

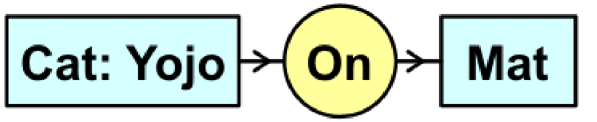
See Peirce’s tutorial on EGs: http://www.jfsowa.com/pubs/egtut.pdf

A Proof by Peirce's Rules



Conclusion: ***Pedro is a farmer who owns and beats a donkey*.**

From EG to CG



Conceptual graphs add more features to the EG foundation:

- More symbols for Boolean operators and quantifiers.
- Point of quantification expressed by a concept box instead of a line.
- Concept box has a type field and a referent field, as in [Cat: Yojo].
- Types are represented by monadic relations or lambda expressions.
- Referent field may have names, indexicals, or generalized quantifiers.

A default ontology designed for mapping to and from NLs:

- Thematic roles and other relations used in linguistics.
- Context boxes for propositions and situations.
- Temporal, modal, causal, and intentional relations.

Operations adapted from DRS and other linguistic systems.

Mapping a Text to CG

A very short story:

> ***At 10:17 UTC, Yojo the cat and a mouse were in the basement of a house. Yojo chased the mouse. He caught the mouse. He ate the head of the mouse.***

Context-free translation:

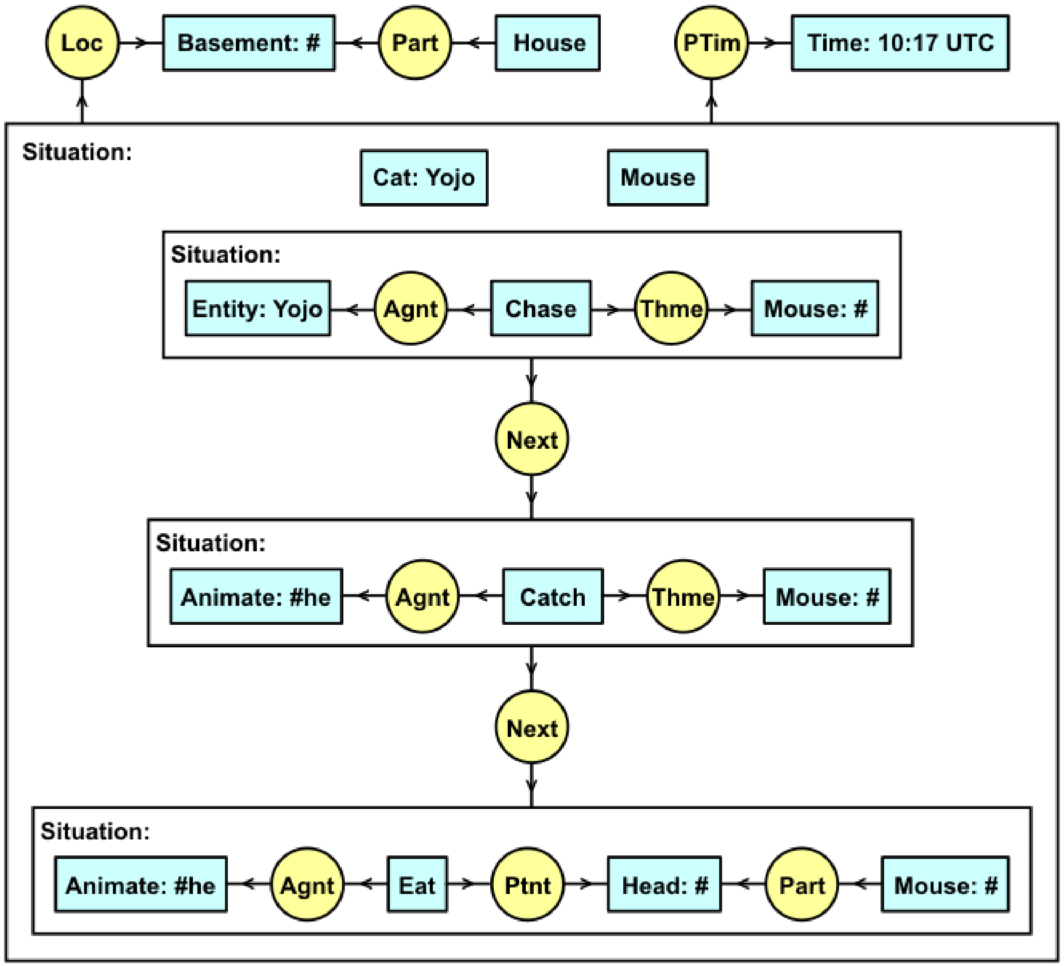
- **Translate each phrase of each sentence without considering the context.**
- **Leave some types and relations underspecified.**
- **Mark unresolved indexicals with the symbol #.**

As more information becomes available from the context,

- **Replace indexical markers with coreference labels.**
- **Specialize (or correct) the type labels on concepts and relations.**

Since DRS contexts are isomorphic to the EG and CG contexts, the DRS rules for indexicals can be adapted to CGs.

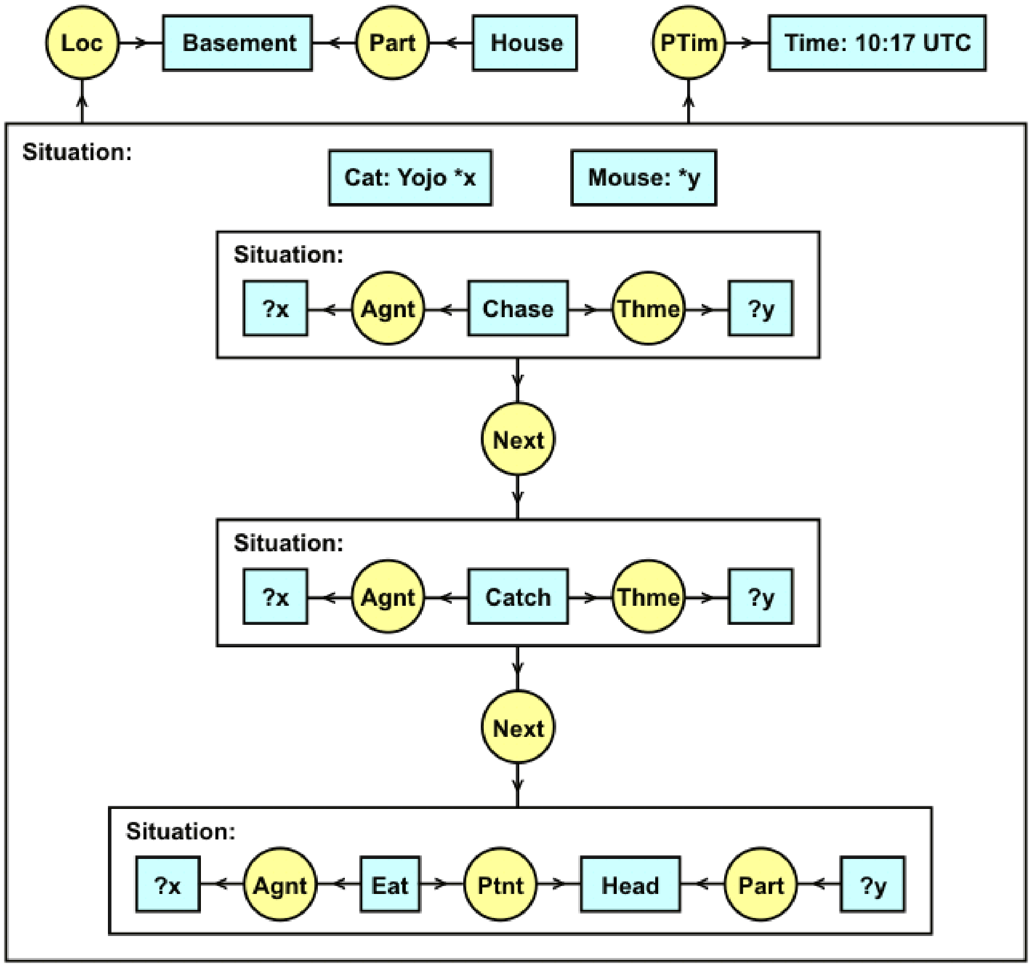
CG with Indexical Markers



The symbol # marks a concept with an unresolved indexical.

Each # marker must be replaced with a coreference label.

CG with Indexicals Resolved



Symbols *x and *y are defining labels; ?x and ?y are bound labels.

In the pure graph notation, coreference is shown by lines.

The Continuum from Scruffy to Neat

DRT and other dynamic logics are neat and formally defined.

They can support useful versions of controlled NLs for narrowly defined language games. *

But nearly all unrestricted speech or writing contains unsolved research problems.

Yet many "scruffy" programs can extract useful information from such language, even though they don't have a formal semantics.

As Einstein and Halmos observed, even mathematicians start with informal language, which they gradually refine.

Any NLP system for unrestricted language must support an open-ended continuum from scruffy to neat language forms.

*** ACE is a system for controlled English that uses DRS: http://attempto.ifi.uzh.ch/site/**

Lattice of Theories

To support the full range of human language, a framework for NLP must be able to represent anything that anyone might ever say.

That framework must include all neat theories, and it must relate them to any kind of language, neat or scruffy.

For any version of logic L, theory X is more general than theory Y, and Y is more specialized than X, if and only if

- X is true of everything (every model for L) for which Y is true.

Generalization determines a *Lindenbaum Lattice* of theories:

- If theory X is more general than Y, write $X \geq Y$ or $Y \leq X$.
- For any X and Y, there is a unique *minimal common generalization*, written $X \cup Y$, such that $X \cup Y \geq X$ and $X \cup Y \geq Y$.
- For any X and Y, there is a unique *maximal common specialization*, written $X \cap Y$, such that $X \cap Y \leq X$ and $X \cap Y \leq Y$.
- The most general theory at the top of the lattice, written $\top$, is true of everything (every model).
- The most specialized theory at the bottom, written $\bot$, is true of nothing.

Model Theory for Mental Models

Every theory in the lattice except the bottom ⊥ has one or more Tarski-style models for which the theory is true:

- **Every model *m* has a set of entities D called the *domain* of *m*.**
- **D has a subset R of *relations* and a subset F of *functions*.**
- **An *atom* is a relation in R or a function in F with an entity in D assigned to each argument position (including the result position of a function).**
- **Some set of atoms is declared to be true.**

Mental models are images of continuous aspects of the world.

They can be described by discrete words or atoms.

But a complete description would specify which atoms are false:

- **Closed world assumption (CWA): Any atom not declared true is false.**
- **Open world assumption (OWA): Some atoms are declared true, some are declared false, and the truth values of the others are unknown.**
- **Dynamic semantics and nonmonotonic reasoning provide systematic methods for filling the gaps in the truth values of OWA models.**

Formalizing Wittgenstein's Satzsysteme

For a given logic (Beweissystem), all the Satzsysteme for that logic can be represented by theories in a lattice.

If the logic is very expressive, the lattice can contain sublattices for many specialized logics.

The universal theory at the top of the lattice contains nothing but tautologies, which are true of every model of every theory.

In Peirce's terms, theories near the top are "sufficiently vague" to characterize a wide range of subjects.

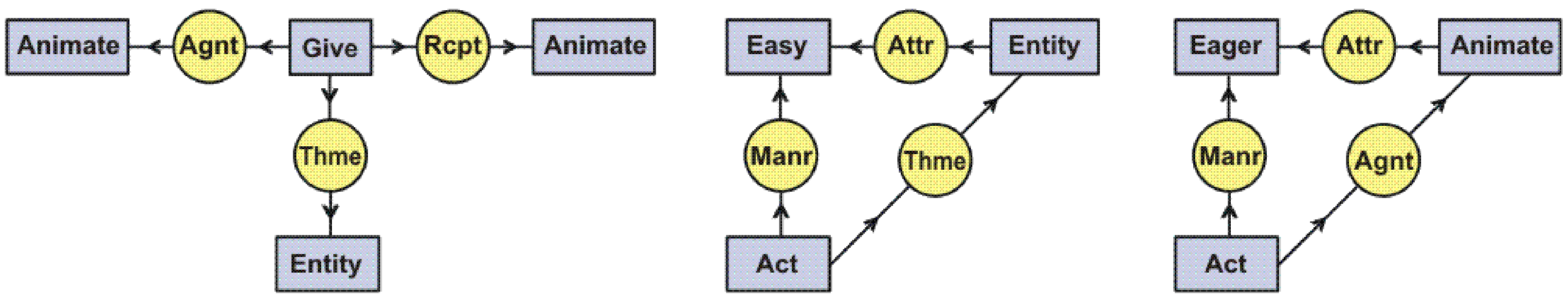
Specialized theories at lower levels are sufficiently "narrow" to be "pretty precise and fairly certain" for more specialized subjects.

But to provide a semantics for language, the theories must be mapped to words, sentences, and extended discourse.

Canonical Graphs

A *canonical graph* is a conceptual graph that represents a pattern or *schema* that is typical for a given concept type.

Canonical graphs for the concept types Give, Easy, and Eager.

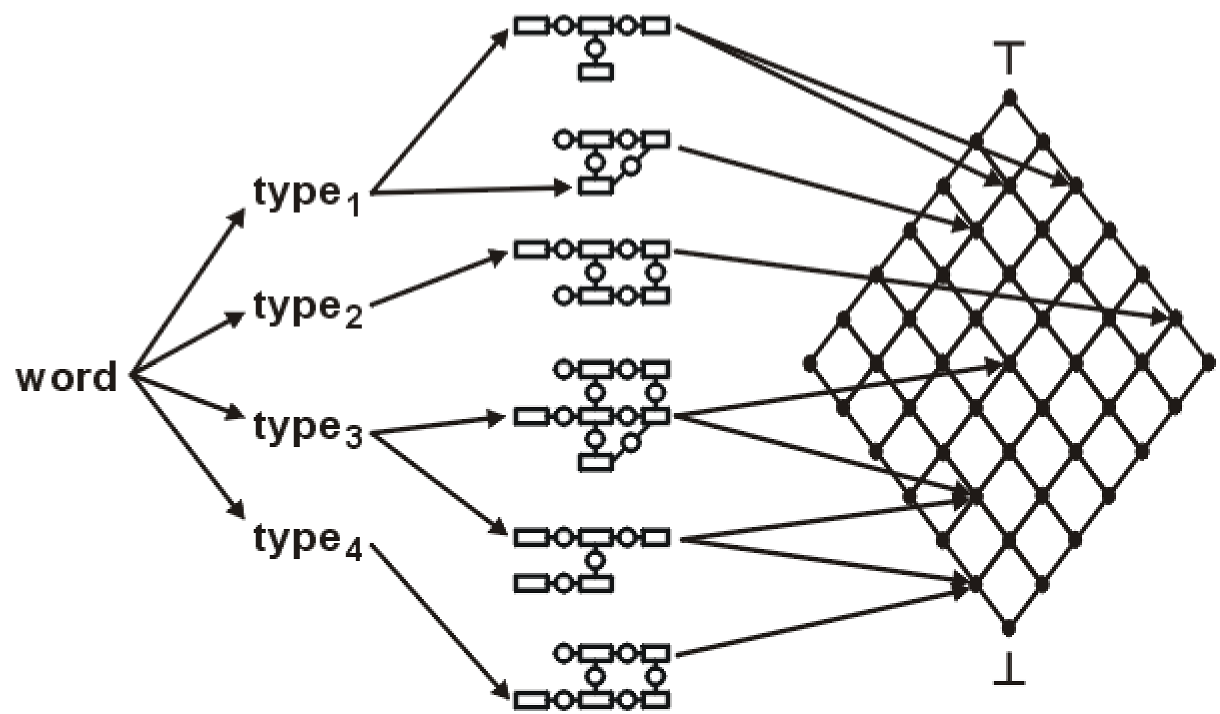


These graphs encode the typical patterns associated with each concept or relation type.

They can be used to select theories or other knowledge relevant to subjects described by those concepts.

The canonical graphs for verbs specify the *case relations* or *thematic roles* and the constraints on concept types. See the IBM-CSLI verb ontology, http://www-csli.stanford.edu/~arunm/

Mapping Word Fans to a Lattice of Theories



words → types → canonical graphs → theories

Supporting Language Games

Three basic mechanisms can map Satzsysteme or language games to and from the sentences of a natural language:

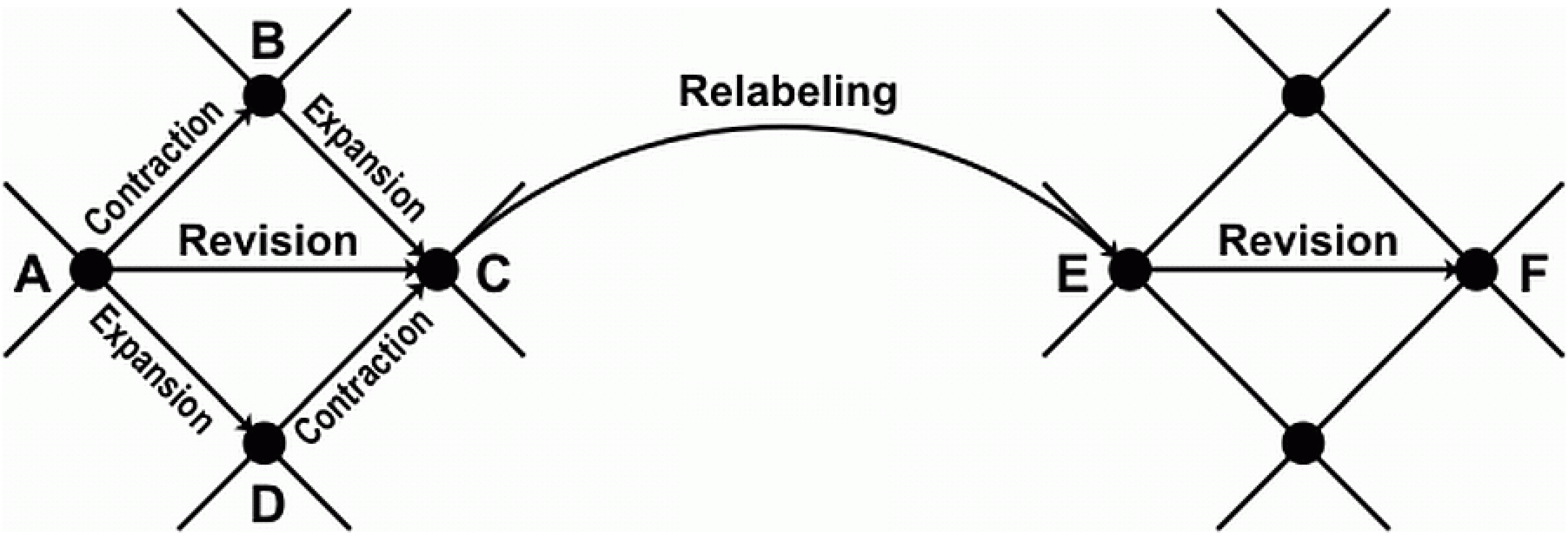
- **Lattice of theories.**
- **Word fans that map lexical items to a hierarchy of concept types.**
- **Canonical graphs that match patterns of words to determine which theory in the lattice is appropriate for interpreting a given sentence.**

But language games require more dynamic operations:

- **A Satzsystem corresponds to a controlled natural language that restricts the syntax, semantics, and vocabulary to a precisely-defined sublanguage.**
- **Language games must be more general than Satzsysteme in order to represent all the dynamic interactions of life.**
- **Language games can grow, change, and be replaced or combined at any time during a conversation.**

Navigating the Lattice of Theories

Methods of belief revision for relating one theory to another:



Four operators: contraction, expansion, revision, and relabeling.

Every method of learning or nonmontonic reasoning determines a strategy for walking or jumping through the lattice of theories.

Learning and Belief Revision

Children learn language by starting with words and patterns of words that are linked to perception and action.

By trial and error, children and adults revise, extend, and adjust their beliefs to make better predictions about the world:

- **Observations generate low-level facts (atoms).**
- **Induction derives general axioms from multiple facts.**
- **A mixture of facts and axioms is an unstructured *knowledge soup*.**
- **Abduction selects some axioms to form a hypothesis (theory).**
- **Analogies relabel a theory of one subject and apply it to another.**
- **Deduction from a theory generates predictions about the world.**
- **Action tests a prediction against reality.**
- **The effects of the action lead to new observations.**

These steps correspond to a walk through a lattice of theories.

Learning a New Theory

Observations generate facts:

```
Tweety is a bird.    Tweety flies.
Daffy is a bird.     Daffy flies.
Hooty is a bird.     Hooty flies.
```

Induction derives general axioms from multiple facts:

```
Every bird flies.
Every flying thing is a bird.
For every x, x is a bird if and only if x flies.
```

Any one of these axioms can be added to a subset of the facts to generate the other facts.

Heuristics give a slight preference for "`Every bird flies.`"

But the other axioms cannot be ruled out.

New Information Triggers Belief Revision

New observation:

`Vampy is not a bird. Vampy flies.`

This observation rules out two options, leaving just one:

`Every bird flies.`

Another observation:

`Tux is a penguin. Tux is a bird. Tux does not fly.`

This observation restricts the universal quantifier:

`Every bird that is not a penguin flies.`

Learning and belief revision can be interpreted as walks through the lattice to find a more appropriate theory.

Proofs in Nonmonotonic Logic

A proof by any method of nonmonotonic logic can be mapped to a walk through a lattice of purely classical FOL theories.

Default logic by Reiter (1980) is a widely used example:

- **Each default theory has two kinds of axioms: classical and default.**
- **Any proof is a sequence of steps: S_0, S_1, …, S_n.**
- **Some steps are classical, and some use a default axiom.**
- **The collection of classical axioms defines some theory C in the lattice.**
- **Each default step S_i makes some assumption A_i in classical FOL.**
- **Adding A_i to the current classical axioms is a revision by expansion.**
- **When the proof ends at step S_n, the classical theory C has been expanded to a more specialized classical theory C′.**
- **In theory C′, the same conclusion can be derived by classical FOL rules.**

Negation as failure (NAF) is a variant of default logic in which any proposition that cannot be proved is assumed to be false.

Dynamic Model Theory

For any first-order logic L, there are two lattices:

- **$\mathcal{T}$ is a Lindenbaum lattice of all theories expressible in L.**
- **$\mathcal{M}$ is the sublattice of $\mathcal{T}$ whose axioms are atoms or negations of atoms.**
- **Every open-world or closed-world model *m* for any theory *t* in $\mathcal{T}$ can be represented by the axioms of some theory in $\mathcal{M}$.**

The lattice $\mathcal{M}$ can describe mental models. Walks through that lattice can describe the dynamic ways they behave and evolve:

- **Every theory *t* other than ⊥ can state constraints that are true of some specialization of every model in at least one sublattice M of $\mathcal{M}$.**
- **A *narrative* is a time-ordered sequence of models along some walk through the models in a sublattice M determined by some theory *t*.**
- **Debates among multiple speakers may cause walks or jumps through the lattice $\mathcal{T}$ that can cause irregular jumps among models in $\mathcal{M}$.**

Modal Logics

For the semantics of modal logics, Kripke's possible worlds can be mapped to the lattice of theories and the lattice of models.

Dunn (1973) showed how with his semantics of laws and facts: *

- **For every Kripke world *w*, there are two theories called the *laws* of *w* and the *facts* of *w*.**
- **The laws of *w* are all propositions necessarily true of *w*.**
- **The facts of *w* are all true propositions about *w* – necessary or contingent.**
- **All the laws of *w* form a theory that is a generalization of all facts of *w*.**
- **The atoms of any model of *w* form a description of the world *w*.**
- **For any worlds w_1 and w_2, the world w_2 is *accessible* from w_1 if and only if the theory of laws of w_1 is a generalization of the theory of facts of w_2.**

The laws of any world can be partitioned into separate theories for different modalities, such as alethic, deontic, or epistemic.

*** See http://www.jfsowa.com/pubs/laws.htm and http://www.jfsowa.com/worlds.pdf**

5. Language Understanding

People relate patterns in language to patterns in mental models.

Simulating exactly what people do is impossible today:

- **Nobody knows the details of how the brain works.**
- **Even with a good theory of the brain, the total amount of detail would overwhelm the fastest supercomputers.**
- **A faithful simulation would also require a detailed model of the body with all its mechanisms of perception, feelings, and action.**

But efficient approximations to human patterns are possible:

- **Graphs can specify good approximations to continuous models.**
- **They can serve as the logical notation for a dynamic model theory.**
- **And they can support a high-speed associative memory.**
- **These features can support formal and informal reasoning.**
- **This combination can also support a semantic interpreter that works in parallel with a parser based on well-known technologies.**

Knowledge Soup

A heterogeneous, loosely linked mixture:

- Fluid, lumpy, and dynamically changing.
- Many lumps are or can be structured in a computable form.
- But they may be inconsistent or incompatible with one another.

In anybody's head, knowledge soup is

- The totality of everything in memory.

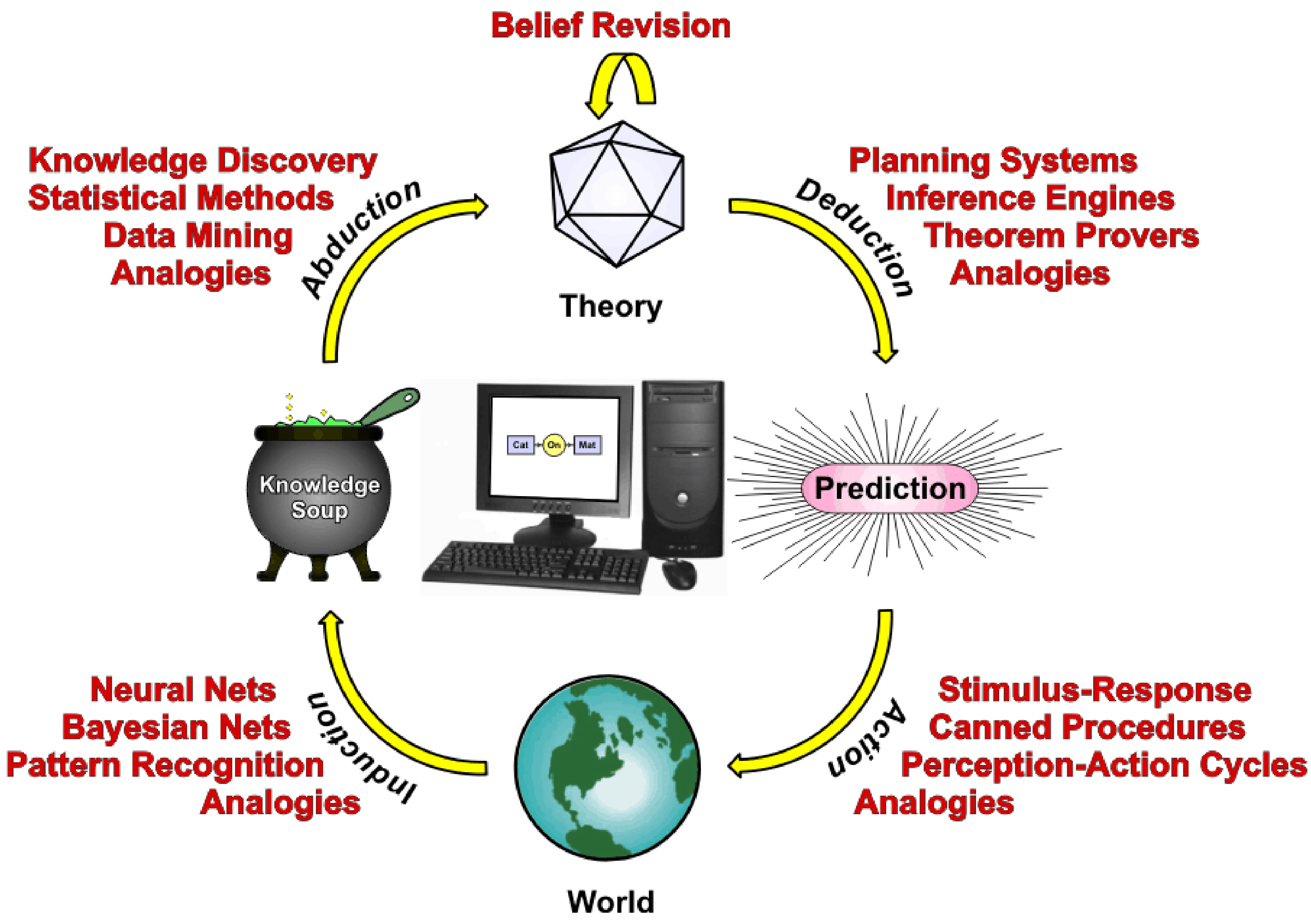
In the WWW, knowledge soup is

- The totality of everything people downloaded from their heads, recorded automatically, or derived by any computable method.

Linked Open Data is good for finding and classifying anything in the soup – whether loose items or structured lumps.

But understanding the contents of the LOD poses the same challenge as understanding natural language.

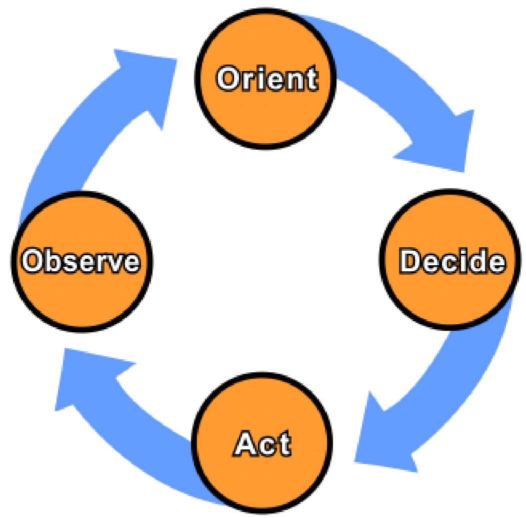
Peirce's Cycle of Pragmatism



This diagram, based on Peirce's writings, shows an integrated view of the dynamic methods for walking through a lattice of theories.

The cycles can be traversed at any speed from seconds to years.

Boyd's OODA Loop



John Boyd drew a four-step diagram for training fighter pilots to observe and respond rapidly.

The first two steps – Observe and Orient – involve the occipital, parietal, and temporal lobes.

The next two steps – Decide and Act – involve the frontal lobes for reasoning and motor control.

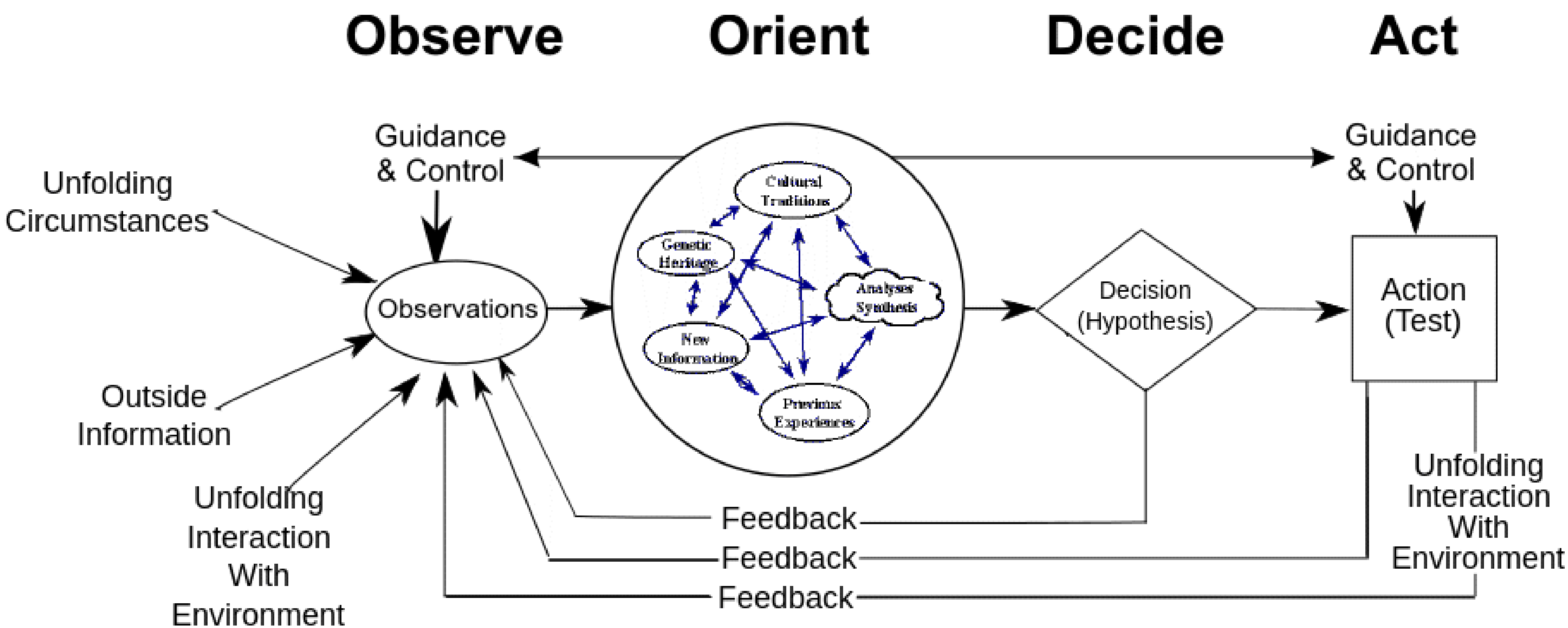
The four steps and the associated brain areas:

1. **Observe: Visual input goes to the primary visual cortex (occipital lobes), but object recognition and naming involves the temporal lobes.**
2. **Orient: Parietal cortex relates vision and touch in "cognitive maps."**
3. **Decide: Reasoning is under the control of the frontal lobes, but other lobes are also involved.**
4. **Act: "Action schemata" are patterns in the premotor cortex. Signals from the primary motor cortex go to the muscles.**

Each step can be traversed in milliseconds for rapid response.

Language and logic would be too slow to control this kind of reasoning.

Extended OODA Loop



Over the years, Boyd added more detail to the OODA Loop.

He applied it to decision-making processes of any kind.

Both versions are consistent with Peirce's cycle.

Diagram adapted from http://en.wikipedia.org/wiki/OODA_loop

Processing the Knowledge Soup

Theories in the lattice are abstractions from the more flexible and vastly more complex connections in the brain.

Structured theories are crystallized from an unstructured soup processed by many different neural mechanisms.

The society of mind by Minsky (1986, 2006):

- There can be no single, unified theory of all neural computation.
- The brain contains many areas specialized for different purposes.
- Learning creates more specializations for every human need.
- The result is a society of interacting heterogeneous modules.
- Knowledge consists of a network of K-lines that connect related aspects in areas of the brain with different specializations.
- Emotions are the driving forces that motivate all operations.

Peirce’s cycle can accommodate a heterogeneous mixture of formal and informal computational modules and methods.

Peirce, Minsky, and Wittgenstein

Peirce's cycle relates the many ways of traversing the lattices:

- Induction, abduction, and deduction are special cases of analogy.
- The cycle may have nested cycles at different levels of granularity.
- Milliseconds: Observe danger, orient, decide, and act (OODA loop).
- Seconds to minutes: Routine analysis, planning, and problem solving.
- Days to years: Complex studies, exploration, and research.

Minsky's society of agents can support multiple paradigms:

- Heterogeneous representations in different areas of the brain.
- Multiple formal and informal computational methods.
- Asynchronous message passing for control and communication.

Wittgenstein's language games support open-ended flexibility:

- Dynamic methods relate the games to perception, action, and reasoning.
- They can be supported by a dynamic model theory.

Applications of VivoMind Technology

The next slide shows a table derived from research reports.

To analyze the reports, an ontology of CGs for chemistry and superconductivity was added to Cognitive Memory.

Then for each report,

- Map each sentence to a conceptual graph.
- Analyze all anaphoric references between the CGs.
- The result is a single, tightly connected CG that links all the CGs for every sentence in the report.
- Store that CG in Cognitive Memory.
- Query Cognitive Memory for the data in each row of the table.
- Store the answers in the table.

In a competition among twelve NLP systems,

- VivoMind got 96% of the entries correct.
- The second best had 73% correct. Two others were slightly above 50%. All the others were below 50%.

Information Extracted from Published Documents

DOE BREMS PROJECT.xlsx

COMPOUND	CURIE TEMP.	SOURCE
Mn3[Cr(CN)6]2 ·16H2O	89 K	A solid-state hybrid density functional theory study
Sr3Ir2O7 in Sr3Ir2O7 single-cr	~ 280 K	Canted antiferromagnetic ground state in Sr3Ir2O7
PrPt2B2C	6 K	Coexistence of superconductivity and magnetic ord
La0:3Nd0:7Pt2:1-	2.8 K	Coexistence of superconductivity and magnetic ord
NdPt2:1B2:4C1:2	3 K	Coexistence of superconductivity and magnetic ord
NdPt1:5Au0:6B2C	3 K	Coexistence of superconductivity and magnetic ord
SmNiC2	= 17.7 K	Commensurate charge-density wave with frustrate
Co0.2Zn0.8Fe2O4. in CdxCo1-	~ 780 K	Does Ti+4 ratio improve the physical properties of C
Zn0.88Co0.12O in ZnO	~ 540 K	Effect of Co doping on the structural; optical and m
La in Sr2-xLaxFeMoO6	425 K	Effect of La doping on the properties of Sr2-xLaxFeM
Fe in Sr2-xLaxFeMoO6	~ 1040 K	Effect of La doping on the properties of Sr2-xLaxFeM
FeSe	~ 305 K	Electronic and magnetic properties of FeSe thin film
Ni-Mn-Ga	= 376 K	Electronic and structural properties of ferromagnet
LaFexSi1 - x13 in La1-zPrz(Fe(	~ 190 K	Enhancement of magnetocaloric effects in La1-zPrz
LaFe0.88Si0.1213 in La1-zPrz(	= 195 K	Enhancement of magnetocaloric effects in La1-zPrz
Co2MnGa in Co2MnGa	600 K	Ferromagnetic resonance in Co2MnGa films with va
HoCrO4 in HoCrO4	17.6 K	Ferromagnetism vs. antiferromagnetism of the dim
Mn3(HCOO)6 in Mn3(HCOO)(	8.0 K	Guest-induced chirality in the ferrimagnetic nanopo
NaZn13- in La0.5Pr0.5(Fe0.88	range from 195 K to 185 K	Large isothermal magnetic entropy change after hy
La2/3Ba1/3MnO3 in La2-3Ba1	range from 300 K to 250 K	Magnetic and neutron diffraction study of La2-3Ba1
CuMnSb in Co1-xCuxMnSb	range from 476 K to 300 K	Magnetic properties of half-metallic semi Heusler C
Nd2 in Nd2-yDyyFe17-xSix	range from 61.46 °C to 236 °	Magnetic properties of iron-rich Nd2-yDyyFe17-xSix
Tb2Fe17 in Nd2-yDyyFe17-xSi	~ 80 °C	Magnetic properties of iron-rich Nd2-yDyyFe17-xSix

Application to Oil and Gas Exploration

Source material:

- **79 documents, ranging in length from 1 page to 50 pages.**
- **Some are reports about oil or gas fields, and others are chapters from a textbook on geology used as background information.**
- **English, as written for human readers (no semantic tagging).**
- **Additional data from relational DBs and other structured sources.**
- **Lexical resources derived from WordNet, CoreLex, IBM-CSLI Verb Ontology, Roget's Thesaurus, and other sources.**
- **An ontology for the oil and gas domain written in controlled English by geologists from the University of Utah.**

Queries:

- **A paragraph that describes a potential oil or gas field.**
- **Analogies compare the query to the documents.**

Answering Questions with VAE

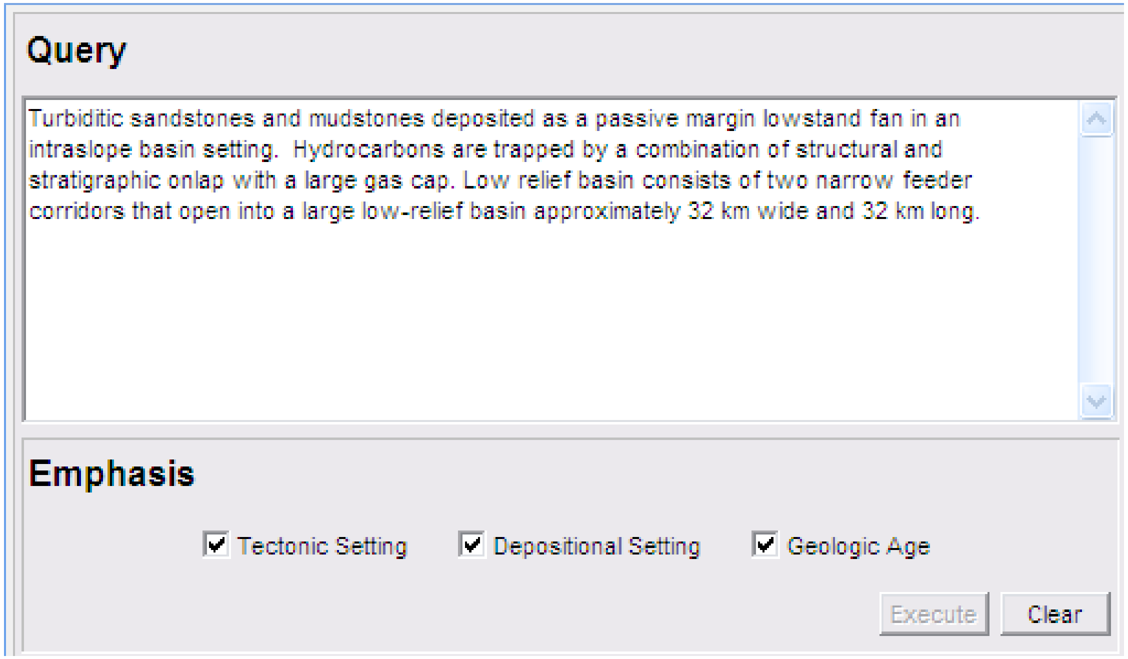
For the sources, either NL documents or structured data:

- Translate the text or data to conceptual graphs.
- Translate all CGs to Cognitive Signatures™ in time proportional to (N log N), where N is the total number of CGs.
- Store each Cognitive Signature in Cognitive Memory™ with a pointer to the original CG and the source from which that CG was derived.
- Use previously translated CGs to help interpret new sentences.

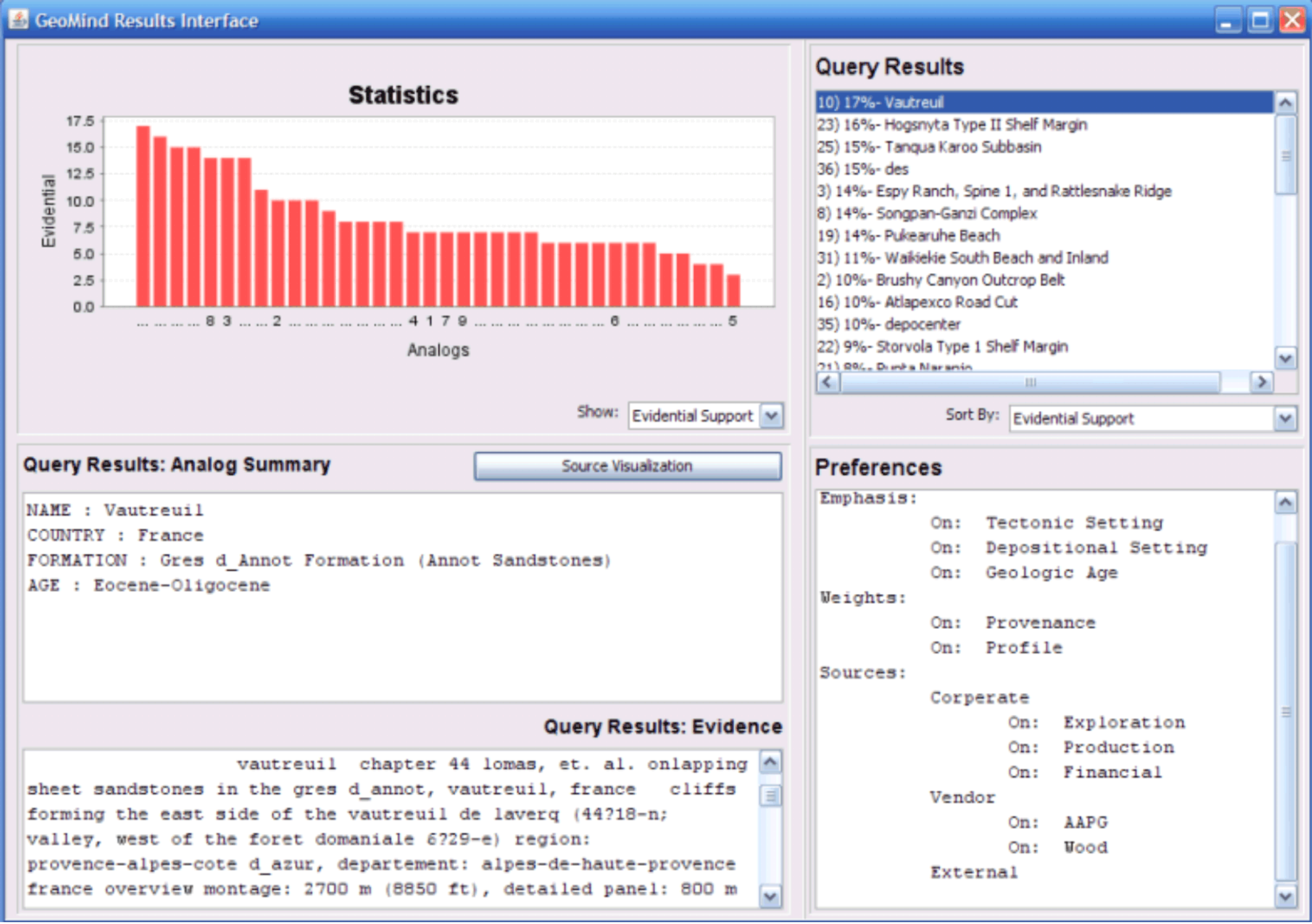
For a query stated as an English sentence or paragraph,

- Translate the query to a conceptual graph.
- Find matching patterns in the source data and rank them in order of semantic distance.
- For each match within a given threshold, use structure mapping to verify which parts of the query CG match the source CG.
- As answer, return the English sentences or paragraphs in the source document that had the closest match to the query.

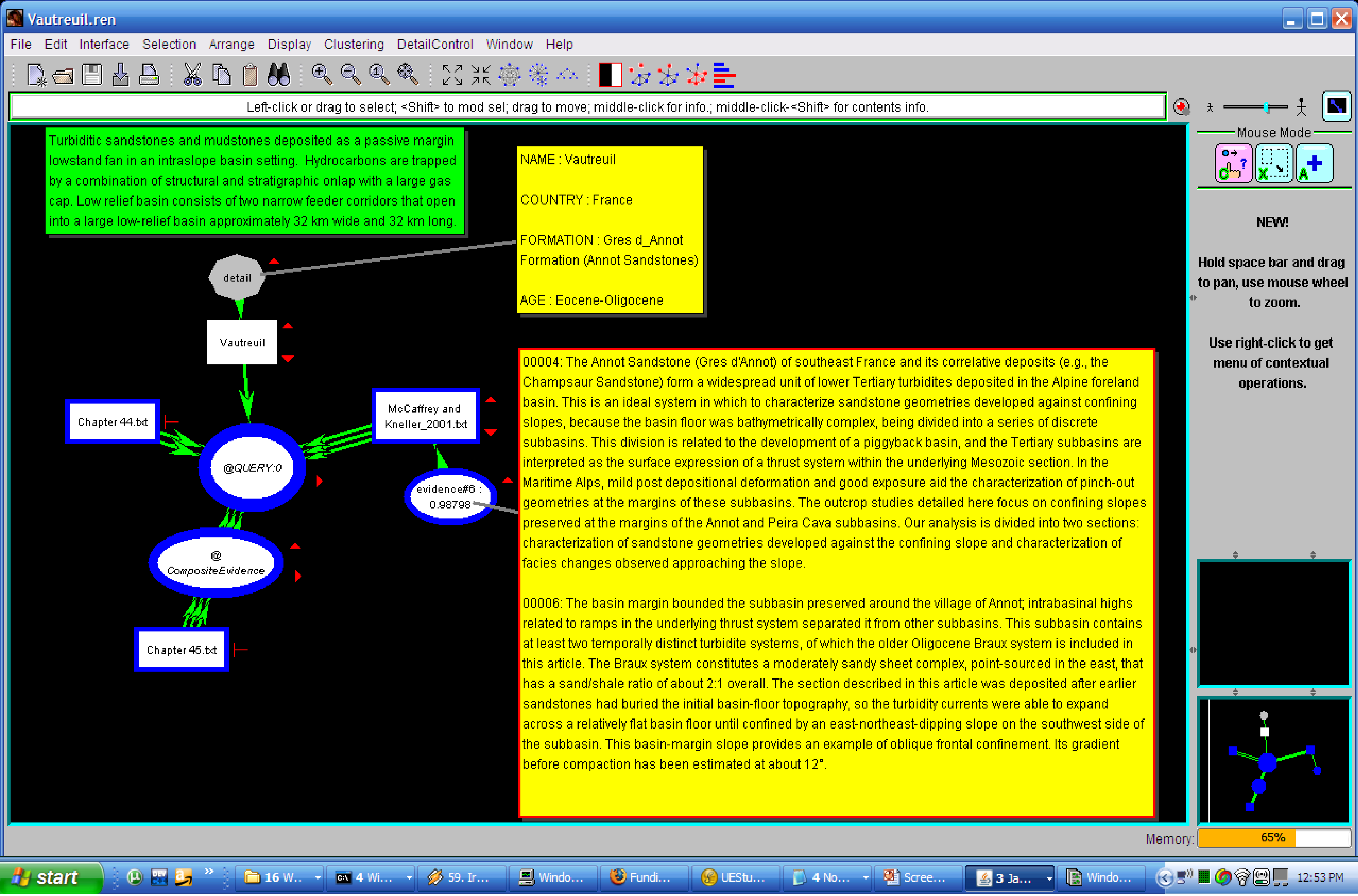
A Query Written by a Geologist



Turbiditic sandstones and mudstones deposited as a passive margin lowstand fan in an intraslope basin setting. Hydrocarbons are trapped by a combination of structural and stratigraphic onlap with a large gas cap. Low relief basin consists of two narrow feeder corridors that open into a large low-relief basin approximately 32 km wide and 32 km long.



Details of the closest matching hydrocarbon fields



Linking the query to the paragraphs that contain the answer

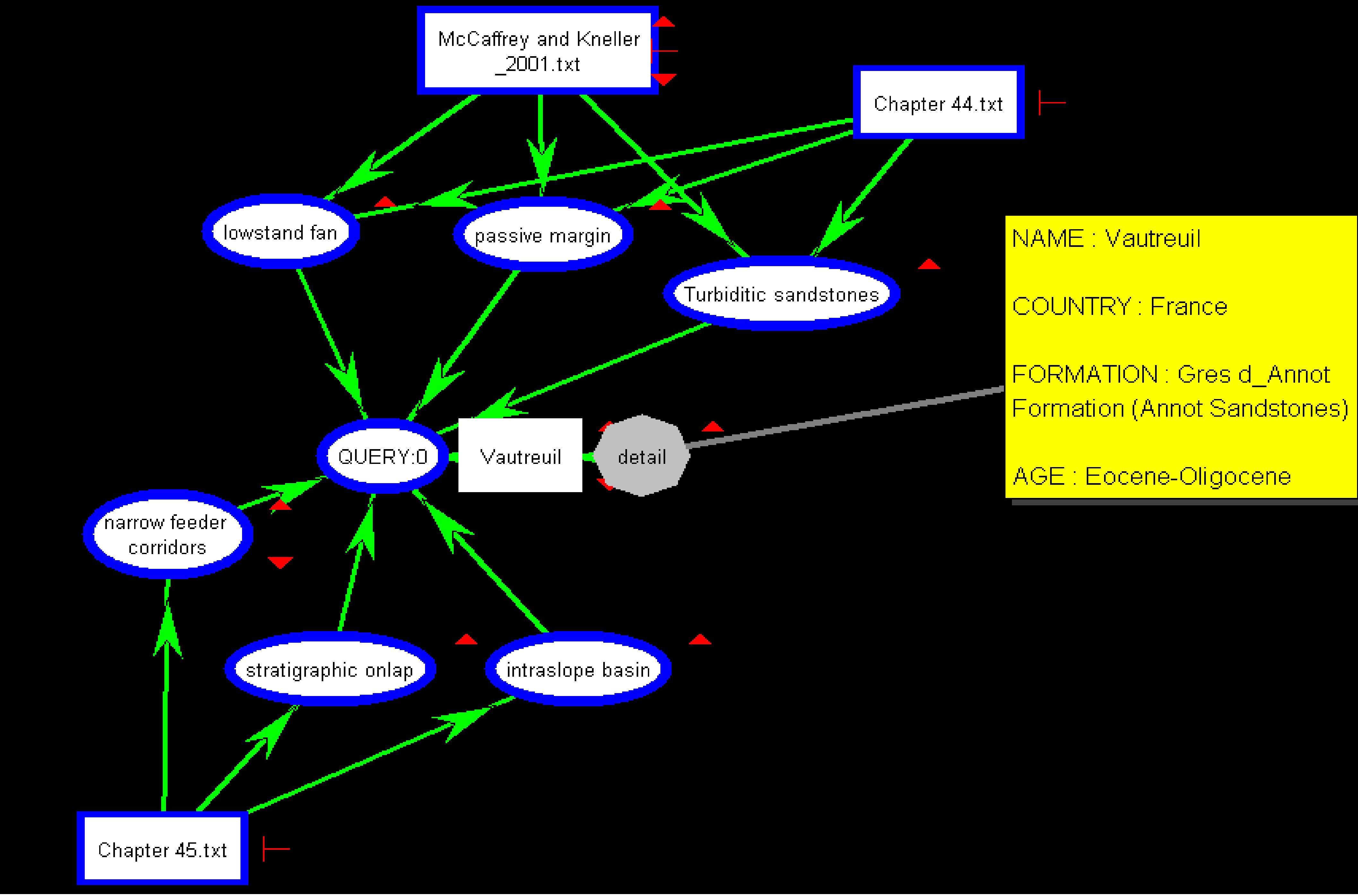
What the Screen Shots Show

Information shown in the previous screen shot:

- The query in the green box describes some oil or gas field.
- The data in the small yellow box describes the Vautreuil field.
- The large yellow box shows the paragraphs in a report by McCarthy and Kneller from which that data was extracted.

The next screen shot shows how the answer was found:

- Many terms in the query were not defined in the ontology: *lowstand fan, passive margin, turbiditic sandstones, narrow feeder cables, stratigraphic onlap, intraslope basin*.
- VLP generated tentative CGs for these phrases and looked in Cogntive Memory to find similar CGs derived from other sources.
- Chapters 44 and 45 of the textbook on geology contained those CGs as subgraphs of larger graphs that had related information.
- Patterns found in the larger graphs helped relate the CGs derived from the query to CGs derived from the report that had the answer.



Using background knowledge from a textbook to find the answer

Emergent Knowledge

When reading the 79 documents,

- VLP translates the sentences and paragraphs to CGs.
- But it does not do any further analysis of the documents.

When a geologist asks a question,

- The VivoMind system may find related phrases in many sources.
- To connect those phrases, it may need to do further searches.
- Some sources can be textbooks with background knowledge that helps VLP interpret the research reports.
- The result consists of conceptual graphs that relate the query to paragraphs in research reports that contain the answer.
- The new CGs can be added to Cognitive Memory for future use.

By a “Socratic” dialog, the geologist can lead the system to explore novel paths and discover unexpected patterns.

Readings

The goal of language understanding,
http://www.jfsowa.com/talks/goal.pdf

Role of Logic and Ontology in Language and Reasoning,
http://www.jfsowa.com/pubs/rolelog.pdf

For more information about VivoMind software:

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

Sowa, John F. (2002) Architectures for intelligent systems, http://www.jfsowa.com/pubs/arch.htm

Sowa, John F., & Arun K. Majumdar (2003) Analogical reasoning, http://www.jfsowa.com/pubs/analog.htm

Sowa, John F. (2003) Laws, facts, and contexts, http://www.jfsowa.com/pubs/laws.htm

Sowa, John F. (2005) The challenge of knowledge soup, http://www.jfsowa.com/pubs/challenge.pdf

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

Sowa, John F. (2011) Future directions for semantic systems, http://www.jfsowa.com/pubs/futures.pdf

For other references, see the combined bibliography for this site:

http://www.jfsowa.com/bib.htm