Dynamic Ontology

A Wittgensteinian Method of Relating Language to the World

John F. Sowa

8 October 2008

Early Optimism about AI

1960: Hao Wang's theorem prover took 7 minutes to prove the first 378 theorems of *Principia Mathematica* on an IBM 704:

Far faster than the two brilliant logicians, Whitehead and Russell.

1960: Emile Delavenay, in a book on machine translation:

"While a great deal remains to be done, it can be stated without hesitation that the essential has already been accomplished."

1965: Irving John Good, in speculations on the future of AI:

"It is more probable than not that, within the twentieth century, an ultraintelligent machine will be built and that it will be the last invention that man need make."

1968: Marvin Minsky, the technical advisor for the movie *2001*:

The HAL 9000 is a "conservative estimate" of the level of artificial intelligence in 2001.

Knowledge is Power

AI reached a level of maturity in the late 1970s and early ’80s:

*** Rule-based expert systems were often faster and more thorough than professionals, and their ability seemed to increase with the number of rules.**

*** New logic-based methods of knowledge representation:**

Logic programming, description logics, discourse representation structures, situation semantics, conceptual graphs, ...

*** Projects to create vast new resources:**

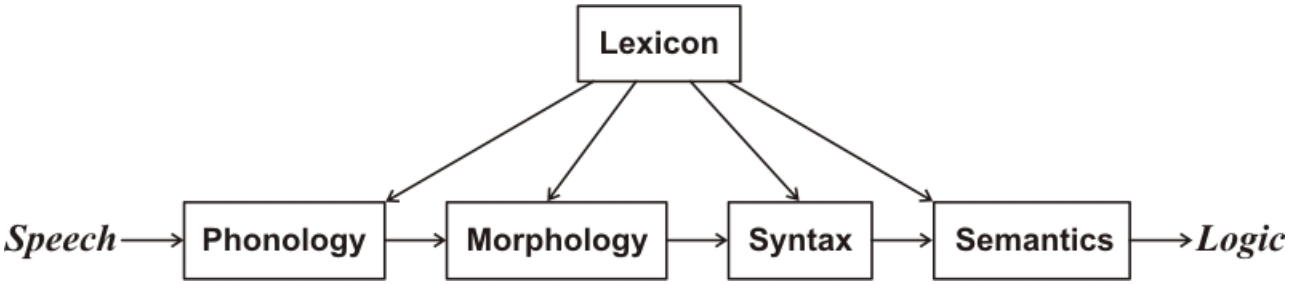
Cyc project by Doug Lenat,

Japanese Electronic Dictionary Research (EDR),

WordNet by George Miller,

...

Classical Natural Language Processing



Frege's principle of compositionality.

Chomsky's distinction between competence and performance.

Tarski-Kripke-Montague's model-theoretic semantics.

A formal ontology for defining word senses.

A linear flow of information from phonology to semantics.

Wittgenstein's First Book

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.

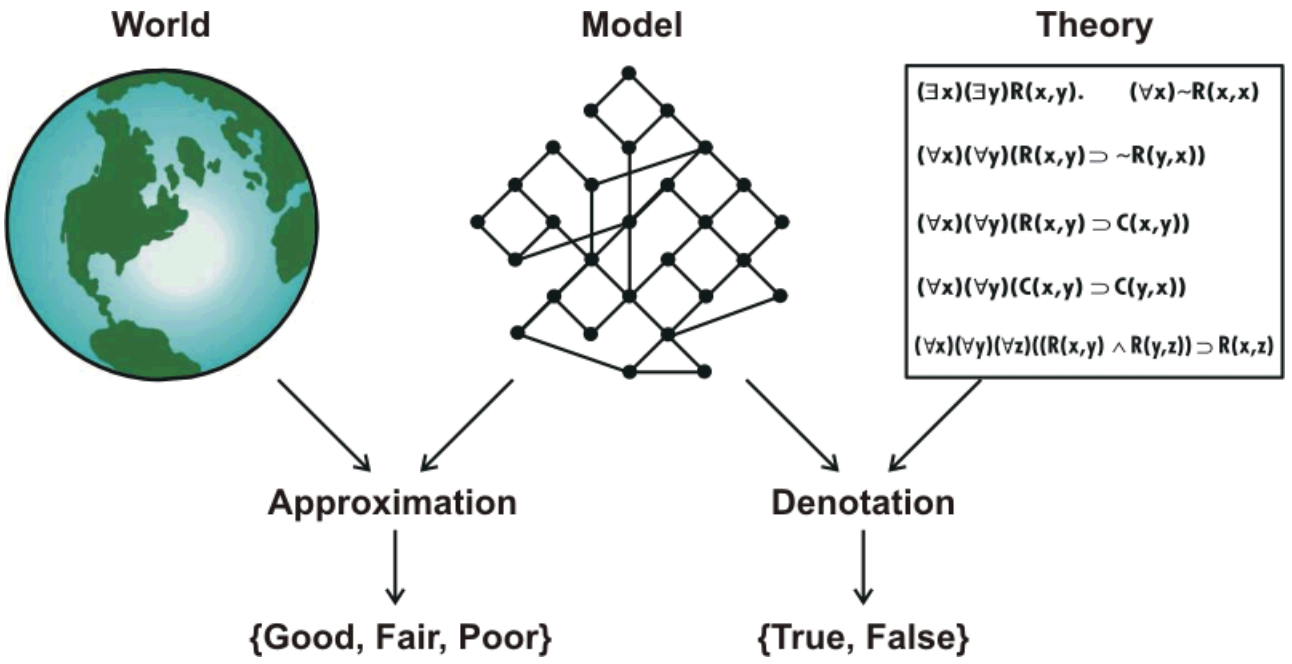
6.13 Logic is not a theory but a reflexion of the world.

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, the HAL 9000 would be ruling the world today.

Model-Theoretic Semantics



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

Since there is exactly one world, there is exactly one model, there is exactly one ontology, and no approximation is conceivable.

Every clear statement is true or false. Everything else is meaningless.

What Kind of Language is Meaningless?

According to the *Tractatus*, everything except scientific statements about the world.

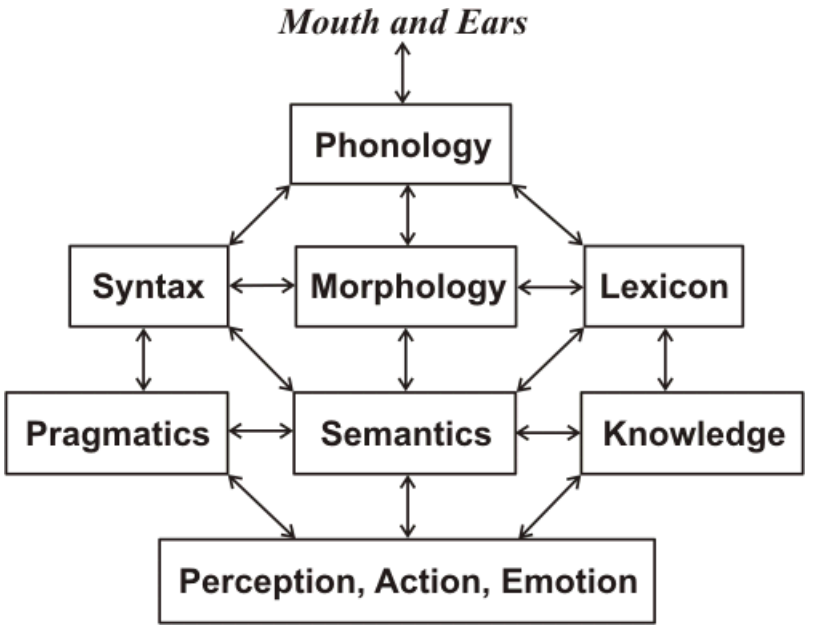
6.421 It is clear that ethics cannot be expressed. Ethics is transcendental. (Ethics and aesthetics are one.)

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.

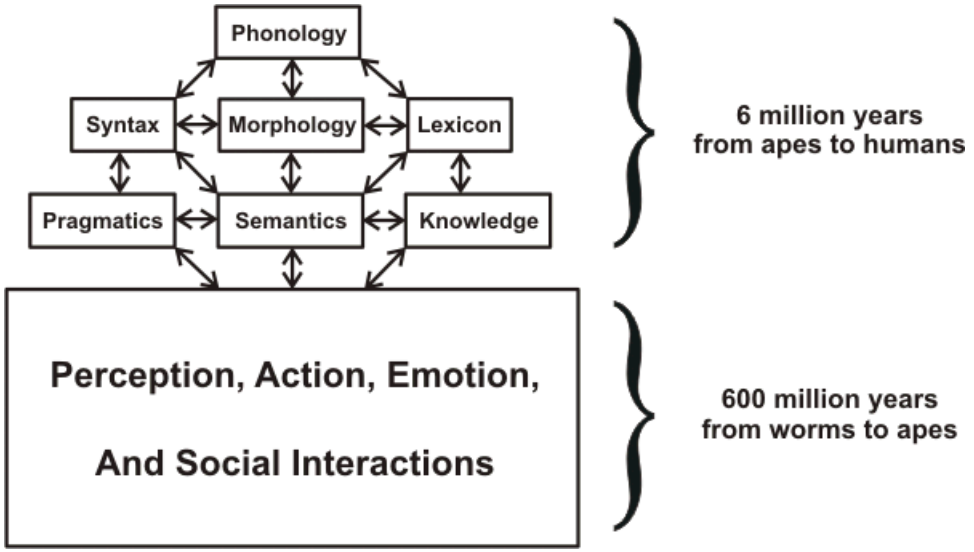
Since the *Tractatus* consists of language about language, it does not state facts about the world. Therefore, it too is meaningless.

Psycholinguistic Studies



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

Evolutionary View of Cognition



All of human cognition except language is at the chimpanzee level.

Language is supported by and integrated with that level.

The Ultimate Understanding Engine

Sentences uttered by a child named Laura at 2 years, 10 months.

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."**

Today's theorem provers are beyond the capability of most humans.

But no computer system today can understand and generate language at the level of a child less than 3 years old.

Yet in terms of the *Tractatus*, most of what Laura says is meaningless.

Limitations of Logical Deduction

2001: Hans Kamp, founder of Discourse Representation Theory,

"The basic concepts of linguistics — and especially those of semantics — have to be thought through anew... Many more distinctions have to be drawn than are dreamt of in current semantic theory."

2004: HALO Project for encoding knowledge from a chemistry text,

Lenat's Cyc system contained two million axioms. Yet neither Cyc nor any other system could read the textbook to acquire knowledge of chemistry.

Despite its large number of predefined axioms, Cyc had no advantage over two other systems with much smaller knowledge bases.

For all three systems, the average cost of translating knowledge from the textbook to axioms in a computable form was $10,000 per page.

Later improvements (better technology and cheaper labor) reduced the cost to about $100 per page.

Statistical Methods

1950s: Information theory, finite-state machines, and Markov models were popular for analyzing natural language texts, but computer resources were inadequate to process large corpora.

1957: Noam Chomsky denounced statistical methods as inadequate.

1980: "Statistics is not AI" was cited as a reason for rejecting papers.

1980s: Fred Jelinek and his group at IBM developed statistical methods that exhausted the resources of IBM mainframes.

1990s: Increased computer size and speed brought statistical methods into the mainstream for learning, machine translation, natural language parsing, and information retrieval.

But statistics alone is not sufficient for language understanding.

Wittgenstein's Transitional Period

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

That analysis led to important innovations:

*** Satzsystem: System of sentences or propositions.**

*** Beweißsystem: 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.

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 Satzsystem, a word is like "a wheel turning idly."**

Identify a Satzsystem with a sublanguage about a very narrow subject that determines the meanings of the words about that subject.

Formalizing Wittgenstein's Satzsysteme

For a given logic, each Satzsystem may be represented by a formal theory that defines the ontology of a narrow subject.

All the theories (Satzsysteme) for a given logic form a lattice.

The universal theory at the top contains nothing but tautologies, which are true in all the theories.

General 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.

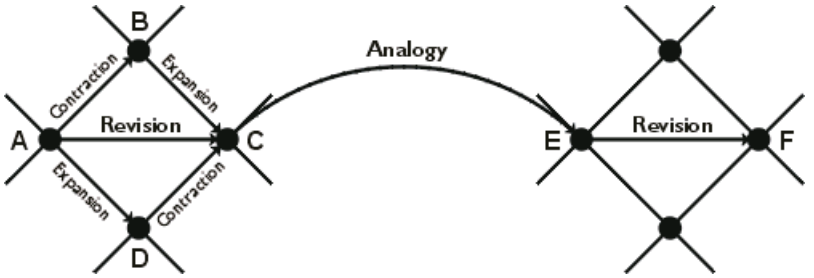
Map the words of a natural language to types of entities and relations.

For each type, represent expected or canonical patterns of usage.

Use those canonical patterns to interpret the semantics of the language.

Navigating the Lattice of Theories

Methods of belief revision for relating one theory or Satzsystem to another.



Many kinds of nonmontonic reasoning may be viewed as strategies for finding a path through the lattice to a preferred theory.

Three AGM operators for belief revision: contraction, expansion, and revision.

Analogy is a fourth method for relabeling concepts and relations.

In learning, new information triggers a search through the lattice to find a new theory that is simpler or more accurate.

Learning a New Theory

Observational data can be stated as a theory with ground-level facts, each of which is independent of the others:

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

Three possible generalizations can reduce the number of axioms:

```
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 these generalizations be added to a subset of the facts to generate the other facts from a more compact basis.

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

But the other options cannot be ruled out.

New Information Causes Belief Revision

Observation:

```
Vampy is not a bird.  Vampy flies.
```

This observation rules out two options, leaving

```
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.
```

The operations of learning and belief revision can be interpreted as walks through the lattice to find a more appropriate theory (or Satzsystem).

Uses and Limitations of Satzsysteme

In his transitional period, Wittgenstein was still addressing the issues about scientific language.

He allowed multiple logics and ontologies, and he relaxed the truth criteria to map entire systems to the world, not just single sentences.

But he did not explicitly accommodate ethics, aesthetics, and metalevel statements that talked about language.

The option of distinguishing the model from the world makes it possible to introduce new features into the model that do not have a direct mapping to anything in the world.

For example, the model might contain abstractions, values, and the words or phrases of a language, independent of any mapping to any observable entities in the world.

Such a generalization would enable language to talk about a broader range of subjects, but it would still be a formal or controlled language, not a truly 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 Beweißsysteme of 1929-30.

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

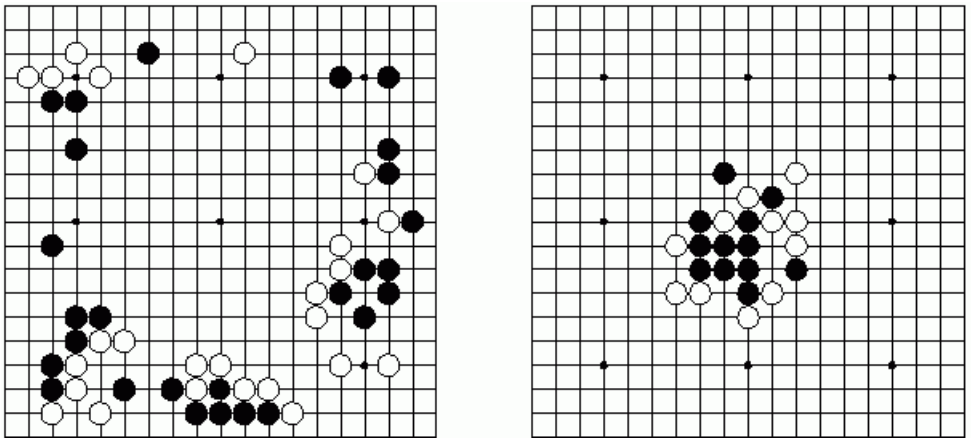
Every use of language is intimately integrated with social activity.

The meaning of every word is grounded in the activity, and it changes with every change in the context and use.

As Wittgenstein said in his notebooks (*Zettel*), language is an "extension of primitive behavior. (For our language game is behavior.)"

As an example, a word is like a piece in the game of chess. Its meaning is derived from the way it is used in the game.

Games of Go and Go-moku



The syntax of making moves in the two games is identical.

But the meaning of any move depends on the goal.

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

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

Two games, same stones, same syntax, but very different patterns.

Examples of Language Games

"And this multiplicity is not something fixed, given once for all; but new types of language, new language-games, as we may say, come into existence, and others become obsolete and get forgotten....

*** Giving orders, and obeying them---**
*** Describing the appearance of an object, or giving its measurements---**
*** Constructing an object from a description (a drawing)---**
*** Reporting an event---**
*** Speculating about an event---**
*** Forming and testing a hypothesis---**
*** Presenting the results of an experiment in tables and diagrams---**
*** Making up a story; and reading it---**
*** Play-acting---**
*** Singing catches---**
*** Guessing riddles---**
*** Making a joke; telling it---**
*** Solving a problem in practical arithmetic---**
*** Translating from one language into another---**
*** Asking, thanking, cursing, greeting, praying."**

Formal Definition of Language Game?

Probably impossible.

Wittgenstein crossed many different academic boundaries:

> **syntax, semantics, pragmatics, logic, ontology, speech acts, scenarios, sublanguage, and genre.**

He emphasized the connection between a language game and social activity:

> **“speaking of language is part of an activity, or of a form of life.”**

But he also compared language games to the different ways of using terms such as “number” in mathematics:

> **“We can get a rough picture of this [the variety of language games] from the changes in mathematics.”**

Implementing Language Games

Wittgenstein's language games have stimulated a lot of debate.

Many logicians considered it "a step in the wrong direction" away from the clarity and logical precision of his first book.

Some computational linguists have found his writings inspirational, but they found it hard to formalize and implement that inspiration.

In fact, Wittgenstein himself would probably object to any attempt to formalize the language games.

The Satzsysteme of his transitional period are more easily formalizable, and they can be considered special cases of language games.

The question is how to support both in a systematic, implementable way.

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).**

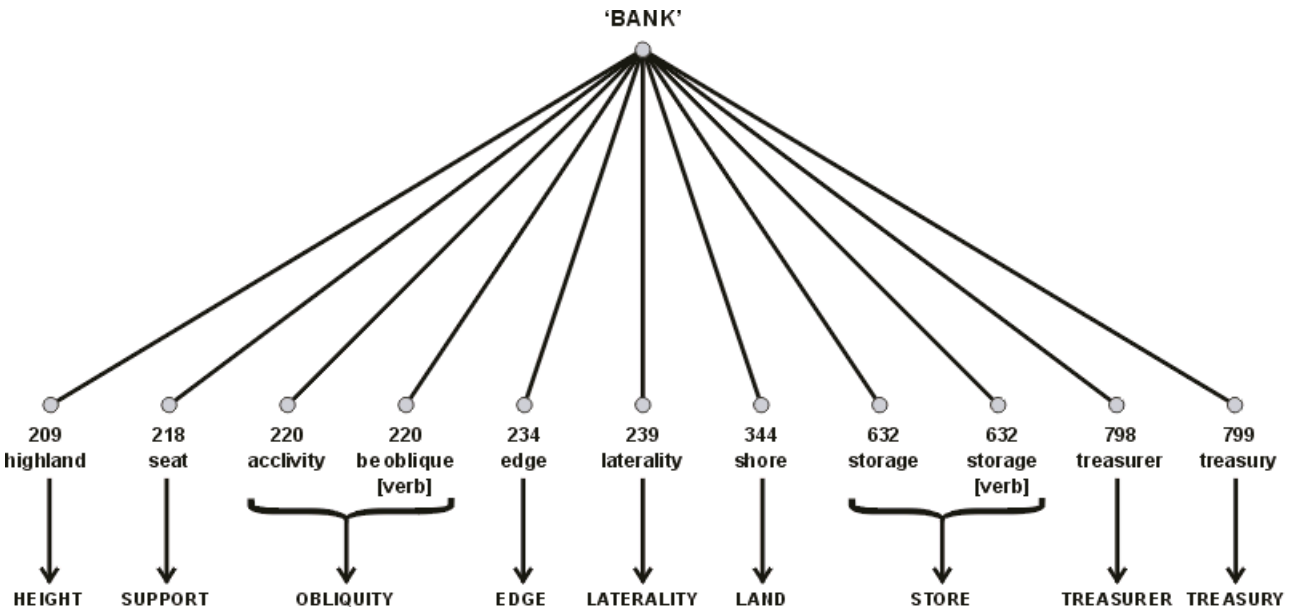
*** Emphasized semantics, not syntax:**

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

*** Developed 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 similar fans.**

*** Combinations of fans “pare down” the ambiguities “by retaining only the spokes that retain ideas which occur in each.”**

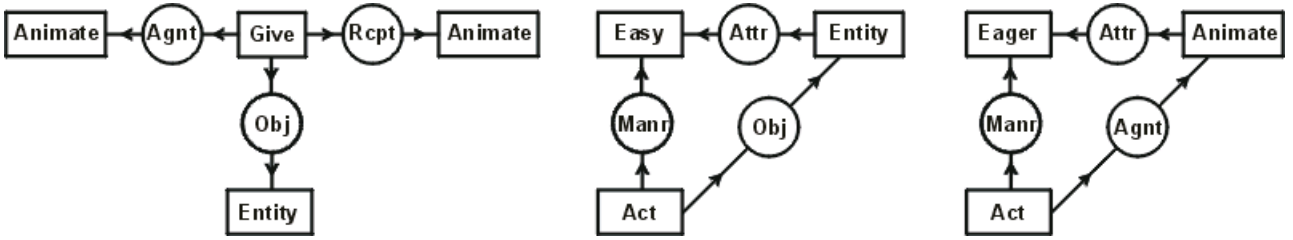
*** For this example,**

— OBLIQUITY 220 is common to 'STEEP' and 'BANK'.

— STORE 632 and TREASURY 799 are common to 'SAVINGS' and 'BANK'.

Canonical Graphs

Conceptual graphs that represent the canonical patterns for a word sense or concept type.



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

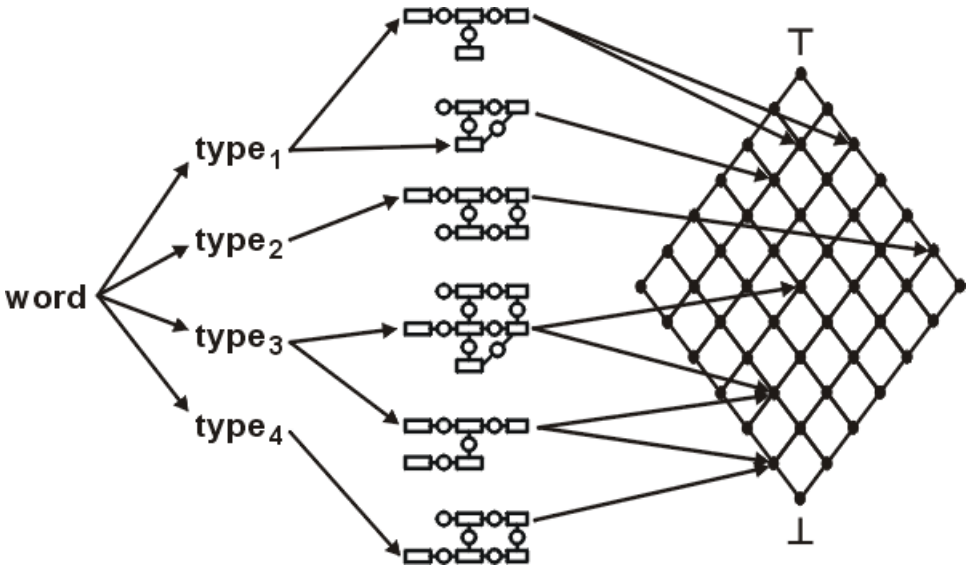
They encode the expected patterns of concepts and relations and select an appropriate theory or Satzsystem for each type or subtype.

But canonical graphs for specialized types may become arbitrarily complex.

E.g., “easy to please person” vs. “easy to read book” vs. “easy to drive car”.

These subtypes of Easy may require very detailed, specialized graphs.

Mapping Word Fans to a Lattice of Theories



words → types → canonical graphs → theories

Supporting Both Versions

Both of Wittgenstein's theories, the Satzsysteme and the language games, can be supported by variants of the same three mechanisms:

1. **Lattice of theories with the belief revision operators.**

2. **Word fans that map lexical items to the type hierarchy.**

3. **Canonical graphs that match patterns of words to determine which theory of the lattice is appropriate to a given phrase, sentence, or paragraph.**

The difference between the Satzsysteme and the language games is in the amount of flexibility and variability they support:

* **A Satzsystem corresponds to a controlled natural language that restricts the syntax, semantics, and vocabulary to a single well-defined sublanguage.**

* **Language games include the Satzsysteme as special cases, but they support much more flexible that allow mixtures of aspects from different games even in different phrases of the same sentence.**

Language and Logic

No version of logic is the foundation for language.

Instead, language is the foundation for all versions of logic.

There is nothing artificial about any version of logic.

They are all perfectly natural language games.

The fundamental method of reasoning is by pattern-matching or analogy.

Induction, deduction, and abduction are disciplined special cases of analogy.

Ontology

A classification and characterization of what exists.

A single, unified, perfect ontology of everything is not possible until after all science has been finished and all research issues have been answered.

Even in that unlikely case, there would still be useful approximations for practical problems:

* **Engineers still use Newtonian mechanics for many purposes, even though relativity and quantum mechanics are known to be "more accurate."**

Aligning the ontologies of multiple agents is as unlikely as using a single ontology for plumbers, bakers, surgeons, bus drivers, police officers, etc.

Each of these groups assume detailed ontologies within their specialties.

They interact successfully when necessary.

But they never align their global ontologies with the other groups.

VivoMind Language Processor (VLP)

A two-level language processor that uses conceptual graphs for both the formal core and open-ended extensions beyond the core.

- *** VLP always generates some conceptual graph for any English sentence.**
- *** But it guarantees a correct interpretation only for sentences stated in Common Logic Controlled English (CLCE).**
- *** For sentences outside CLCE, VLP will generate a translation to CLCE as an echo that the user can verify.**

VLP is currently under development, but it is being designed as a successor to two earlier systems:

1. **Intellitex, which was used for processing language by the VivoMind Analogy Engine (VAE).**
2. **A conventional syntax-directed translator of CLCE to logic.**