Concept Mapping

John F. Sowa

VivoMind Intelligence, Inc.

Abstract. The task of knowledge representation has two parts: the first is to analyze some body of knowledge and identify the relevant concepts, relations, and assumptions; the second is to translate the result of the analysis into some notation that can be processed by computer. Neither part is easy, but the first is far more difficult. Natural languages are capable of expressing anything that can be stated in any artificial language with the same level of detail and precision, but they can tolerate any degree of vagueness during the process of analysis. Artificial languages, such as the many variants of symbolic logic, are valuable because they do not tolerate vagueness, but what they say so precisely may have no relationship to what the author intended. The various notations for logic are designed to represent the final precise stage, but they provide no intermediate forms that can bridge the gap between an initial vague idea and its ultimate formalization. Natural languages can represent every stage from the most vague to the most precise, but no version of fuzzy logic or related variants can come close to the flexibility of natural languages.

The vagueness is not caused by natural language, but by the fact that people seldom have a clear idea of what they want to say before the analysis has been completed. Engineers have a pithy characterization of the phenomenon: “Customers never know what they want until they see what they get.” Plato's dialogs illustrate the kind of analysis that is required. Similar dialogs are necessary when programmers or engineers analyze a vague wish list (also called a *requirements document*) in order to generate a formal specification. Those dialogs always take place in natural languages, often supplemented with hastily scribbled diagrams, but not in any version of logic, fuzzy or precise.

This talk discusses a range of representations from informal to formal and compares four notations that are being used in various stages of knowledge acquisition, analysis, and representation: the informal Concept Maps, the semiformalized Topic Maps, the formal Conceptual Graphs, and the formal, but highly readable Common Logic Controlled English (CLCE). These and other similar notations have found useful niches in the process of analysis and representation, but it is important to recognize their different characteristics and areas of applicability.

The following slides were presented in the track on Technology, Instruction, Cognition and Learning (TICL) at the AERA Conference, San Francisco, 10 April 2006.

The Problem of Knowledge Representation

As stated by the logician 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.

And by the poet 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.

Poetry and logic are complementary approaches to a common problem: developing patterns of symbols that capture important aspects of life in a memorable form.

Concept Maps

A technique for analyzing and diagramming relationships.

- Classic reference: Novak & Gowin (1984) *Learning How to Learn*.
- Informal, easy to learn, and easy to use.
- Surprisingly effective as a tool for eliciting knowledge from domain experts who may have no training in any formal notation.
- Basis for R & D projects at the Institute for Human-Machine Cognition (IHMC).
- IHMC has developed guidelines for drawing “propositionally coherent” Concept Maps.
- With appropriate constraints and conventions, Cmaps have been used as a highly readable front end to formalized notations, such as RDF and OWL.

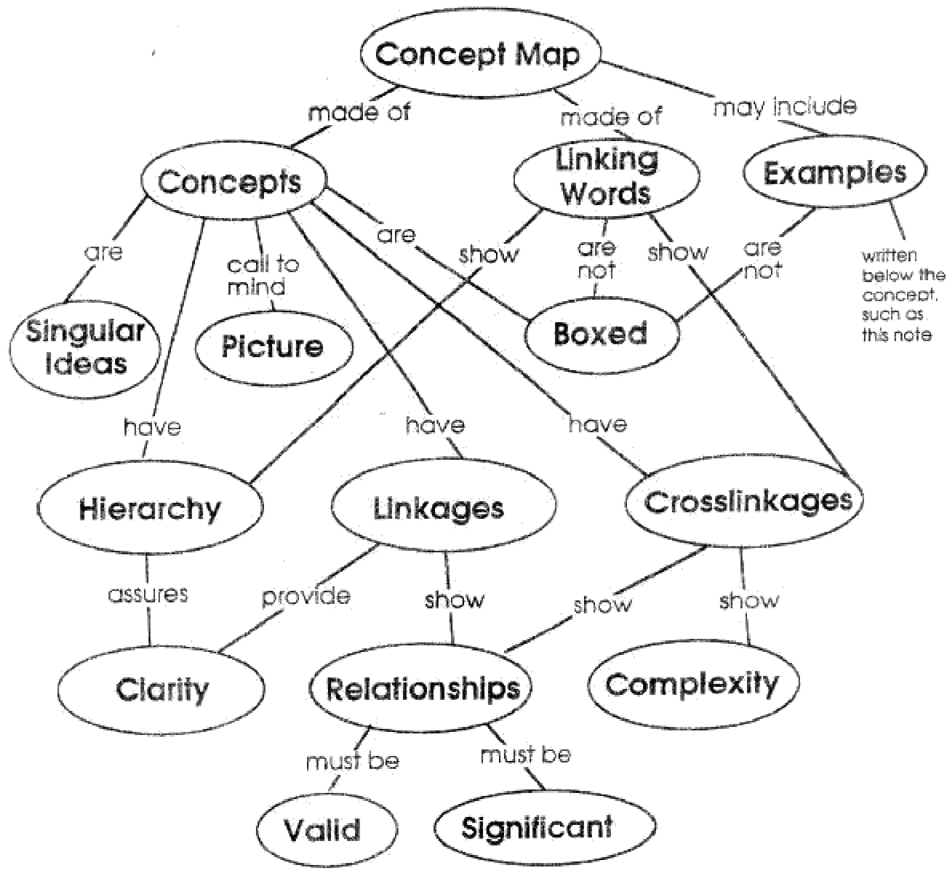
For further information about Cmaps and software for designing and using them, see

http://cmap.ihmc.us/

For a Cmap that contains recommendations for designing good Cmaps, see

http://65.212.118.148/servlet/SBReadResourceServlet?rid=1064009710027_279131382_27088&partName=htmltext

A Concept Map About Concept Maps



Source: http://www.mc.maricopa.edu/dept/d43/glg/Study_Aids/concept_maps/conceptmaps1.gif

---

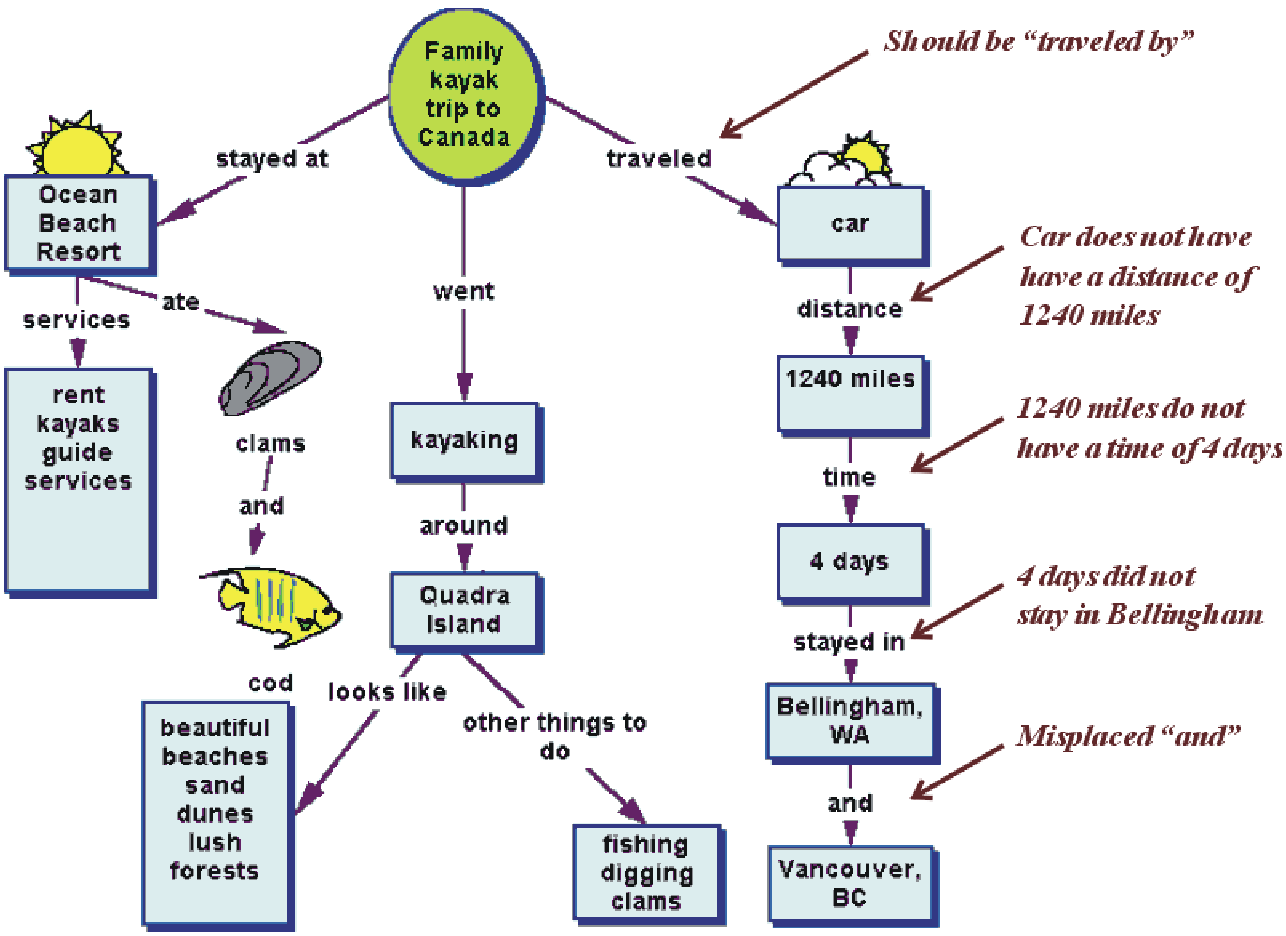
Problems with Cmaps

Informality makes them very easy to learn.

But undisciplined use can create Cmaps that are even more ambiguous than an English sentence:

- Typical error: Mapping a sentence fragment to a path of several nodes.
- Worse error: Ignoring implicit relations expressed by English syntax.
- Sophisticated errors: Trying to express features of logic that cannot be expressed in a simple network, such as quantifiers and Boolean relations other than "*and*".

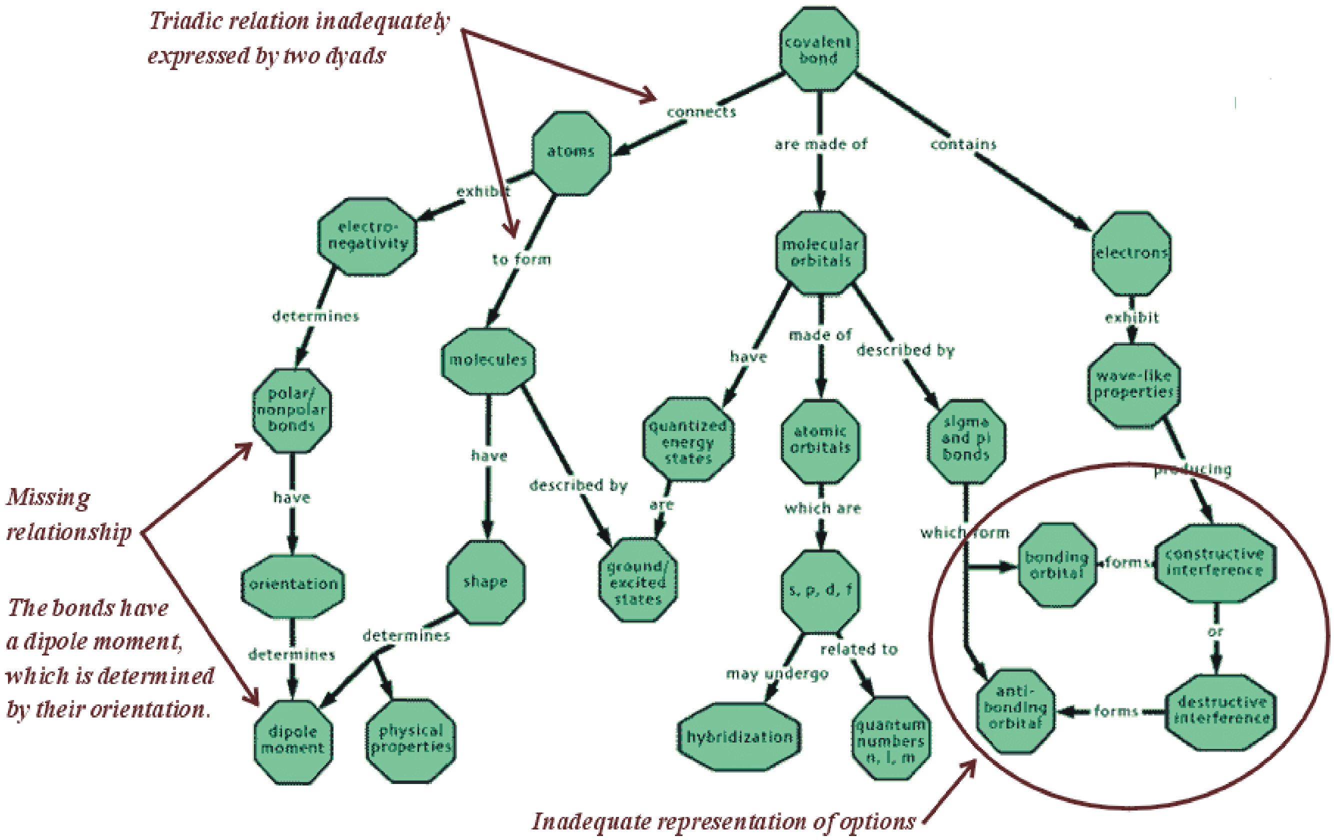
Some Typical Errors



Source: http://www.graphic.org/desmap.html

This Cmap was used as an illustration of the drawing technology and the available graphics. Unfortunately, it is a rather poor illustration of the use of Cmaps to convey information, either to people or to computer systems. It violates the guidelines for a "propositionally coherent" Cmap.

More Sophisticated Problems



Source: http://qsad.bu.edu/ed/TeacherLog/conceptgifs/covalent.gif

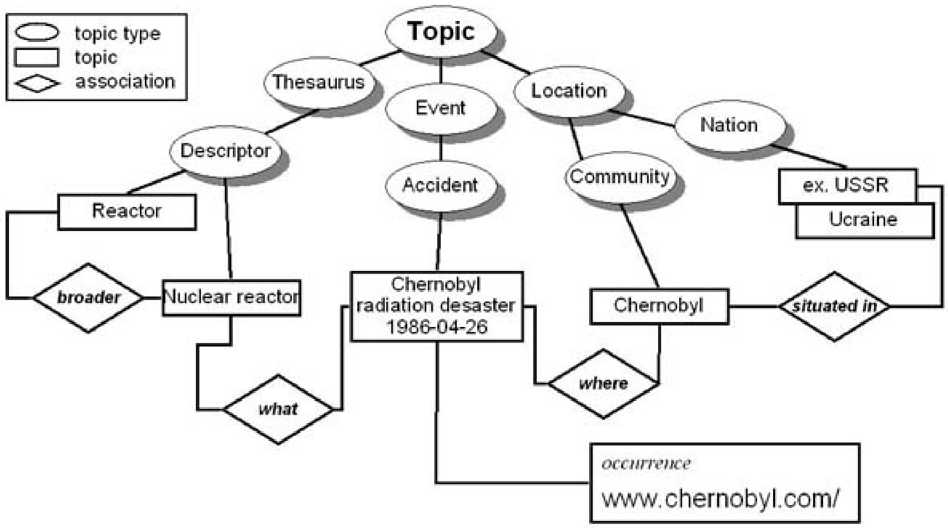
---

Topic Maps

Another graph notation for representing relationships.

- Topic Maps have systematic conventions for representing roles and relationships, which are similar to the guidelines for propositionally coherent Cmaps.
- ISO standard for XTM — an XML-based representation Topic Maps.
- Many guidelines and handbooks with conventions and sample ontologies for using Topic Maps in various domains.
- The XTM standard defines a formal syntax for Topic Maps.
- But there is no formal semantics associated with the XTM syntax.

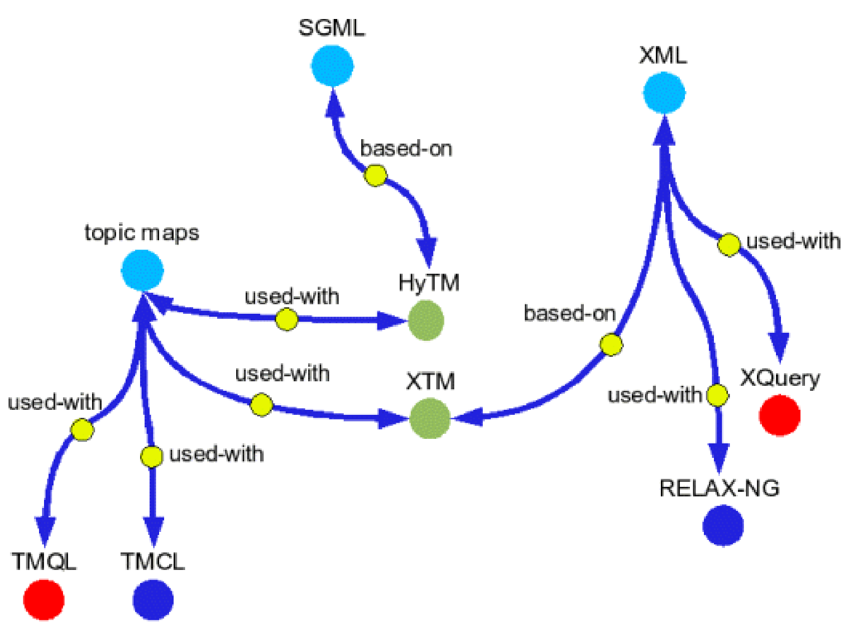
Topic Map for the Chernobyl Disaster



Source: http://www.idealliance.org/papers/xmle02/dx_xmle02/papers/03-05-03/03-05-03.html

The XTM standard formalizes the syntax, but there is no model theory to formalize the semantics. This topic map, for example, mixes metalevel information about the type hierarchy of events and locations, a different kind of metalevel information about the thesaurus that contains terms, and instance-level information about an event that occurred at a particular time and place.

Topic Map for XTM



Relational Graphs

A flat network notation for logic with no provision for grouping propositions by nesting or partitioning.

- Examples: Concept Maps, Topic Maps, RDF, OWL, Type Hierarchies, Entity-Relationship Diagrams, frame systems, and many semantic networks.
- Can be used to express information at the instance level, the metalevel, or both.
- But cannot express the full range of quantifiers and Boolean operators of first-order logic.
- Not general enough to express if-then rules, SQL queries and constraints, or declarative sentences in English.

Many of the errors in using Cmaps result from attempts to go beyond the expressive power of relational graphs.

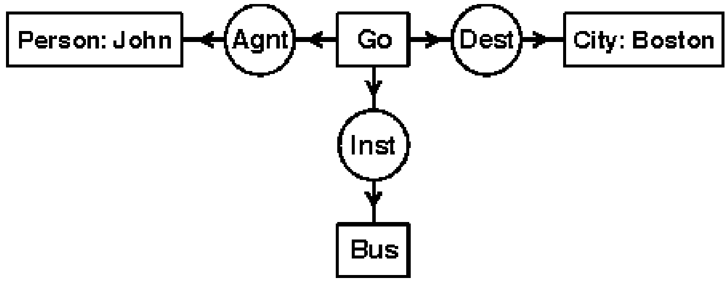
Conceptual Graphs

A notation for logic based on the Existential Graphs of Charles Sanders Peirce and the semantic networks of AI.

- Goal: Express the propositional content of natural language texts.
- Extend relational graphs by adding quantifiers and boxes for enclosing propositions that may be negated or linked by various Boolean and metalevel operators.
- First-order subset of CGs standardized by the draft ISO standard for Common Logic.
- With a proposed extension for metalanguage, Common Logic in any dialect, including CGs, can support the Cyc Language.
- The relational subset of CGs can be used to represent the same kinds of relational graphs expressed by Concept Maps or Topic Maps.
- Mapping unrestricted Concept Maps to CGs, however, would require tools that could resolve ambiguities and add information that is not expressed in the Cmaps. Such tools would normally require sophisticated reasoning ability or some human assistance.

Conceptual graphs are usually easier to read than the equivalent in predicate calculus or CLIF notation, but drawing them requires just as much sensitivity to the nuances of language and logic.

A Simple Conceptual Graph

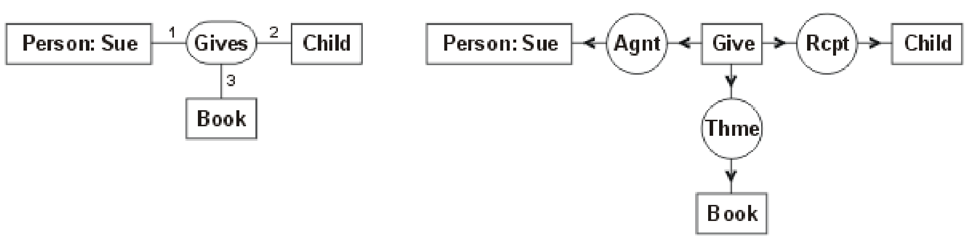


John is going to Boston by bus.

This CG uses only the relational subset of conceptual graphs. With the addition of boxes for representing negated contexts, CGs can represent full first-order logic. By supporting arbitrary relations attached to context boxes, CGs can support metalevels, modality, and intentionality.

Choice of Relations

Two choices of relations:



Sue gives a child a book.

The conceptual graph on the left uses a triadic conceptual relation of type **`Gives`** with one argument for Sue, the second argument for a child, and the third argument for a book. The CG on the right uses a concept of type **`Give`** with three attached conceptual relations: the agent relation (with type label **`Agnt`**) shows that Sue is the giver; the recipient relation (type **`Rcpt`**) shows that the child receives the book; and the theme relation (type **`Thme`**) shows that the book is the thing that is given. These dyadic relations, which are usually called *thematic relations* or *case roles*, are convenient for translations to and from natural languages, but any CG that uses one set of conventions can be translated to or from a CG that uses a different set by means of equivalence rules, also stated in CGs.

Common Logic

Syntax and model-theoretic semantics defined by the draft ISO standard: http://cl.tamu.edu

Three logically equivalent notations in the standard: Common Logic Interchange Format (CLIF), Conceptual Graph Interchange Format (CGIF), and an XML-based notation for Common Logic (XCL).

CGIF for the CG that represents the sentence *John is going to Boston by bus*.

```
[Go *x] [Person: John] [City: Boston] [Bus *y]
(Agnt ?x John) (Dest ?x Boston) (Inst ?x ?y)
```

Equivalent in CLIF:

```
(exists ((x Go) (y Bus))
   (and (Person John) (city Boston)
        (Agnt x John) (Dest x Boston) (Inst x y) ))
```

Equivalent in Common Logic Controlled English (CLCE):

```
The person John goes to the city Boston by a bus.
```

Following is the CGIF for the sentence *Sue gives a child a book*, which corresponds to the conceptual graph with a single triadic relation:

```
[Person: Sue] [Child *x] [Book *y] (Gives Sue ?x ?y)
```

Following is the CGIF for the corresponding CG with three dyadic relations;

```
[Person: Sue] [Child *x] [Book *y] [Gives *z]
(Agnt ?z Sue) (Rcpt ?z ?x) (Thme ?z ?y)
```

And following is the equivalence rule for mapping one form to the other:

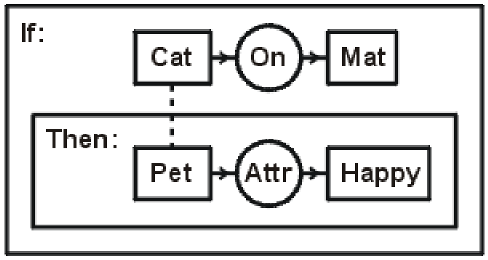
```
[@every *x1] [@every *x2] [@every *x3]
[Equivalence: [Iff: (Gives ?x1 ?x2 ?x3)]
     [Iff: [Give *x] (Agnt ?x ?x1) (Rcpt ?x ?x2) (Thme ?x ?x3)] ]
```

CGIF statements can always be translated to CLIF statements and vice versa, but some forms such as the CG with a triadic **`Gives`** relation may have to be translated to a CG with the dyadic relations that express the linguistic cases before they can be translated to CLCE.

CLCE is not in the standard, but tools are being developed to translate CLCE to and from other CL dialects. For some discussion of controlled natural languages, see

http://www.jfsowa.com/logic/ace.htm

Beyond Relational Graphs



CGIF:

```
[If: [Cat *x] [Mat *y] (On ?x ?y)
   [Then: [Pet ?x] [Happy *z] (Attr ?x ?z) ]]
```

CLIF:

```
(not (exists ((x Cat) (y Mat)) (and (On x y)
   (not (exists ((z Happy)) (Attr x z))) )))
```

CLCE:

```
If a cat is on a mat, then the cat is a happy pet.
```

The context boxes with type labels **`If`** and **`Then`** are defined in terms of negated propositions, as the translation to CLIF demonstrates. Following is an example of CGIF, CLIF, and CLCE for representing mathematical statements:

CGIF:

```
[If: (SetOf [Cat *x1] [Dog *x2] [Elephant *x3] | [Set *x])
    [Then: (IsIn ?x1 ?x) (IsIn ?x2 ?x) (IsIn ?x3 ?x)]]
```

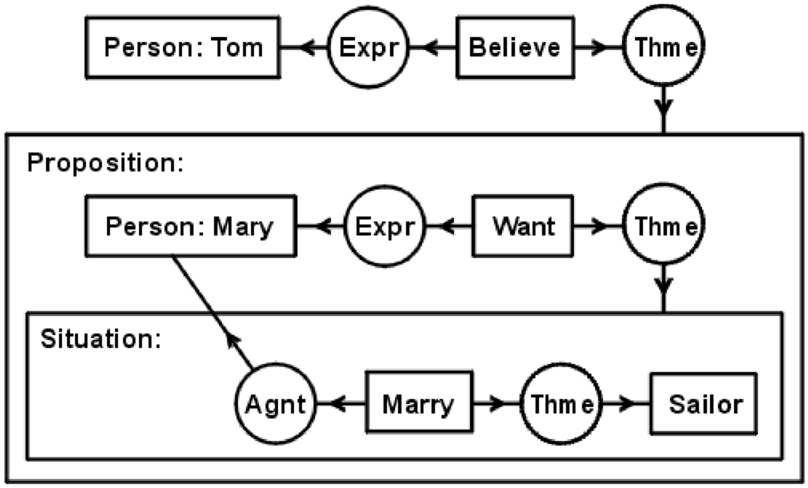
CLIF:

```
(not (exists((x Set)(x1 Cat)(x2 Dog)(x3 Elephant)) (and (= x (SetOf x1 x2 x3))
    (not (and (IsIn x1 x) (IsIn x2 x) (IsIn x3 x))) )))
```

CLCE:

```
If a set x is the set of (a cat, a dog, and an elephant),
then the cat is in x, the dog is in x, and the elephant is in x.
```

Proposed Extension to Common Logic



This CG has contexts of type **`Proposition`** and **`Situation`**, which are the themes of the *intensional* verbs *believe* and *want*. A proposed extension to Common Logic, which is not yet being considered for the ISO standard, would permit metalevel statements about the nested propositions. In CLIF, the keyword **`that`** is used to mark a proposition about which some metalevel statement is being made. The same keyword is used in CLCE.

CGIF:

```
[Person: Tom] [Believe: *x] (Expr ?x Tom)
(Thme ?x [Proposition:
   [Person: Mary] [Want: *z] (Expr ?z Mary)
   (Thme ?z [Situation:
      [Marry: *w] (Expr ?w Mary) (Thme ?u [Sailor]) ])])
```

CLIF:

```
(exists ((x Believe)) (and (Person Tom) (Expr x Tom)
   (Thme x (that (exists ((z Want)) (and (Person Mary) (Expr z Mary)
      (exists ((u Situation)) (and (Thme z w)
         (Dscr u (that (exists ((w Marry) (v Sailor))
            (and (Expr w Mary) (Thme u v)) )))))))))
```

CLCE:

```
Tom believes that Mary wants a situation
   such that Mary marries a sailor.
```

Note that what Tom believes is a proposition, but what Mary wants is a situation. When a CG box of type **`Situation`** encloses a proposition, the description relation **Dscr** is used in CLIF to show that the situation is being described by the proposition.

Conclusions

As Whitehead said, "the problem is to discriminate exactly what we know vaguely."

Writing a precise definition of any aspect of knowledge, even in one's own native language, is not easy.

Flexible notations, such as Cmaps, have proved to be useful, but many steps are needed to convert them to a formal specification.

Most people can easily read controlled English, such as CLCE, but they require considerable training before they can write it.

Conceptual graphs are an intermediate notation that can be used in tools that simplify the mappings to and from the various notations.

But designing good tools with good human factors is still a major research project.

Related Readings

For a discussion of the problems of knowledge representation,

> The Challenge of Knowledge Soup
> http://www.jfsowa.com/pubs/soup.htm

For the relationship of ontology to logic, metadata, metalanguages, and semiotics,

> Ontology, Metadata, and Semiotics
> http://www.jfsowa.com/ontology/ontometa.htm

Model-theoretic foundation of logics with multiple metalevels and nested contexts:

> Laws, facts, and contexts: Foundations for multimodal reasoning
> http://www.jfsowa.com/pubs/laws.htm

Draft ISO standard for Common Logic:

> Home page for the CL draft and ongoing discussions
> http://cl.tamu.edu

For the quotations by Whitehead and Frost,

> Whitehead, Alfred North (1937) "Analysis of Meaning," *Philosophical Review*, reprinted in A. N. Whitehead, *Essays in Science and Philosophy*, Philosophical Library, New York, pp. 122-131.

> Frost, Robert (1963) *A Lover's Quarrel with the World* (film), WGBH Educational Foundation, Boston.