# CONCEPTUAL LEXICON FOR CROSS LANGUAGE COMMUNICATION

**S. Rajendran**

Amrita University, Coimbatore
*rajushush@gmail.com*

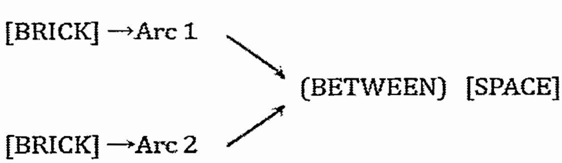
## 1. Introduction

The proposed conceptual lexicon has concepts as its entries. The concepts are independent of a specific language. The meanings of concepts are given in terms of conceptual graphs. From conceptual graphs the surface representation of lexical items belonging to a particular language can be derived. Conceptual graphs emphasize semantics. The earliest forms called existential graphs were invented by the philosopher Chrales Sander Peirce (1897) as a graphical notation for symbolic logic. Lucien Teniere (1959) used similar graphs for his dependency grammar. The earliest form implemented on a computer was the correlational nets by Silvio Cecato (1961), who used them as intermediate language for machine translation. There is philosophical and psychological evidence that conceptual graphs are mental representations unbounded by knowledge of a particular language.

The proposed lexicon can be manipulated to generate a text in the form of a target language. The theory propounded by Sowa (1984) has been exploited to suit our purpose. The four levels of representations proposed for the generative lexicon (Pustejovsky, 1995 Pustejovsky and Boguraev. 1993) are taken into account. The semantic representations in WordNet (Pike Vassion, 1999a,b) are also kept in mind while writing the meaning of a lexical item by means of conceptual graphs.

## 2. Conceptual graphs

Concepts are language independent ones derived from percepts. A conceptual graph is a finite, connected, bipartite graph. The two kinds of nodes of the bipartite graph are concepts and conceptual relations. Every conceptual relation has one or more arcs, each of which must be linked to some concept. For example, a space between a brick and a brick can be represented as follows (Sowa 1984: 72):




Conceptual graphs form a knowledge representation of language based on linguistics, psychology, and philosophy. In the graphs, concept nodes represent entities, attributes, states, and events, and relation nodes show how the concepts are interconnected. Distinctions can be made between simple and complex concepts. Simple concepts are basic concepts from which complex concepts can be derived. Conceptual graphs form a knowledge representation of language based on linguistics, psychology, and philosophy.

### 3. Semantic network

The concept types CAT and TOMATO map directly to percepts. Other types like PRICE, FUNCTION and JUSTICE have no sensory correlates. Abstract concepts acquire their meanings not through direct associations with percepts, but through vast networks of relationships that ultimately link them to concrete concepts. A conceptual graph has no meaning in isolation. The description of the concept, MAN is represented as follows:

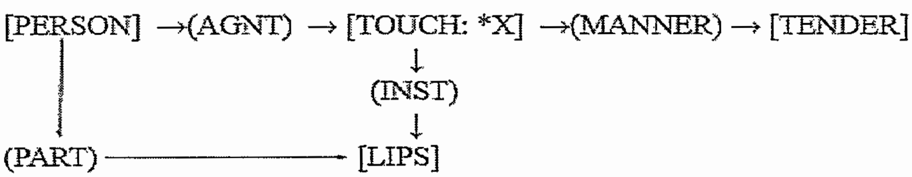
[MAN] → (ISA) → [HUMAN BEING] → (ISA) → [ANIMAL]

### 4. Abstraction and definition

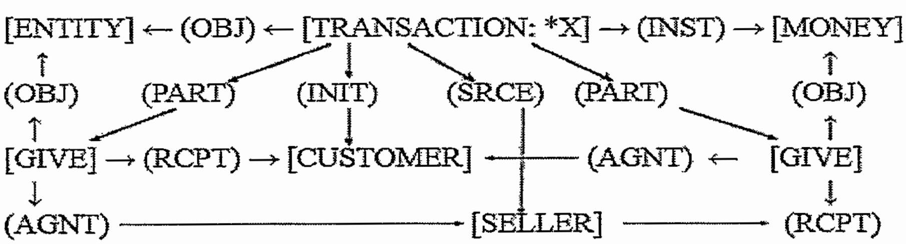
Definition can specify a type in two different ways: i) by stating necessary and sufficient conditions for the type, and ii) by giving a few examples and saying that everything similar to these belong to the concerned type. 'Conceptual graphs' support the type definitions by genus and differentia as well as through schemata and prototypes.

Type definition for KISS (Sowa, 1984: 106)

type KISS(x) is



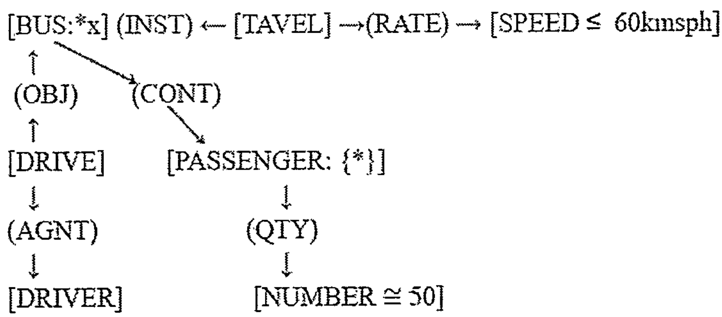

Type definition of BUY (x) is (Sowa, 1984: 110)




### 5. Schemata

Schemata are similar in structure to type definition. Yet concept type may have at most one definition, but arbitrarily many schemata. Type definitions present the narrow notion of a concept, and schemata present the broad notion. Type definitions are obligatory conditions that state only the essential properties, but schemata are optional defaults that state the commonly associated accidental properties. Schemata show the typical ways in which a concept may be used, but they do not describe a typical instance of a concept.

Schema for BUS (x) is (Sowa, 1984: 129)




## 6. Prototype

A prototype is a typical instance. Instead of describing a specific individual, it describes a typical "average individual". A Schema for ELEPHANT might specify a range of characteristics for elephants or a range of behaviours and habitats for elephants. A prototype ELEPHANT would combine and restrict such schema to describe a typical elephant.

Proto type for ELEPHANT (x) is

[ELEPHANT:*x] –

(CHAC) → [HEIGHT: @ 3.3 m]

(CHAC) → [WEIGHT: @ 5400 kg]

(COLR) → [DARK-GREY]

(PART) → [NOSE]-

(ATTR) → [PREHENSILE]

(IDNT) → [TRUNK]

(PART) → [EAR] {*}] –

(QTY) → [NUMBER:2]

(ATTR) → [FLOPPY]

(PART) ? [TUSK: {*}] –

(QTY) → [NUMBER:2]

(MATR) → [IVORY]

(PART) ? [LEG: {*}]

(QTY) → [NUMBER:4]

(STAT) → [LIVE] –

(LOC) → [CONTINENT: {Africa|Asia}]

(DUR) → [TIME: @ 50 YEARS]

## 7. Conceptual representation

Some of the conceptual relations listed in Sowa (1984: 415) are adopted to suit our purpose.

agent. (AGNT) links [ACT] to [ANIMATE], where the ANIMATE concept represents the actor of the action.

attribute. (ATTR) links [ENTITY:*x] to [ENTITY:*y] where *x has an attribute *y.

cause. (CASE) links [STATE:*x] to [STATE:*y] where *x has a cause *y.

characteristic. (CHRC) links [ENTITY:*x] to [ENTITY:*y] where *x has a characteristic *y.

destination. (DEST) links [ACT] to [ENTITY] towards which the action is directed.

experience (EXPR) links [STATE] to [ANIMATE], who is experiencing that state.

instrument. (INST) links [ENTITY] to [ACT] in which the entity is causally involved.

source. (SRCE) links an [ACT] to an [ENTITY] from which it originates.

## 8. Lexical relations

WordNet is an online lexical resource which is built over an ontology in which the building blocks are synsets (i.e. synonymy sets) and the synsets are connected to each other by a network of relations, both semantic and lexical relations. The important ones of these relations are: Hyponymy-Hypernymy, Meronymy-Holonymy, Troponymy and Entailment. These relations can also be made use of when we prepare a conceptual dictionary for cross-language representation. The following examples depict the above mentioned relations.

[ANIMAL] → (HYPER)→ [MAMMAL]

[COW] → (HYPO) → [MAMMAL]

[TABLE] → (HOLO) → [LEG]

[PROFESSOR] →(MEMB) → [DEPARTMENT]

[WHEEL] → (MERO) →[CART]

[WALK] → (TROPO)→ [LIMP]

[SNORE] →(ENTA)→ [SLEEP]

## 9. Sample of conceptual lexicon

The following is the sample of the conceptual lexicon. Each item is a concept and the concepts will be mapped against the lexical items in a language.

[CAT] → (ISA) → [ANIMAL]

[PENCIL] → (ISA) → [INSTRUMENT]

↓

(FUNCT)

↓

[WRITING]

The verbs will be provided with their argument structures. A frame of arguments will be given with their necessary relations. The verbal concept ACT will be represented in the following fashion.

[ACT] → (ISA) → [EVENT]-

(AGENT) → [ANIMATE ENTITY]

[ARRIVE] → (ISA) → [EVENT] -

(AGENT) → [MOBILE-ENTITY]

(GOAL) → [LOCATION]

## 10. Lexical and conceptual structures

Each natural language has a well-organized lexical and syntactic system. Each domain of knowledge has a well-organized conceptual system. Complexities arise because each language tends to use and reuse the same words and lexical patterns in many different conceptual domains.

The lexical structures are

a) Relatively domain independent

b) Dependent on syntax and word forms

c) Highly language dependent

And the conceptual structures are

a) Highly domain dependent

b) Independent of syntax and word forms

c) Language independent, but possibly culture dependent

When there are cross-linguistic similarities in lexical patterns, they usually result from underlying conceptual similarities. In English the verb *give*, for example, takes a subject, object, and indirect object. Other languages may have different cases marked by different prepositions, postpositions, inflections, and word order; but the verb which roughly has the same meaning as *give* also has three participants – a giver, a thing given, and a recipient. In all languages, the three participants in the conceptual pattern lead to three arguments in the lexical patterns.

The distinction between lexical structures and conceptual structures addresses the following things:

a) Lexical structures are oriented towards language. The representation developed here is strongly influenced by linguistic theories of syntax and thematic roles.

b) Conceptual structures are designed for representing knowledge about the world. They may grow too large to be expressed in a single sentence, and they may contain concepts types that cannot be expressed by a single word.

c) Since they can be represented by similar structures, the same operations can be used on them. Furthermore, lexical structures can be converted to deeper conceptual structures by a step-by-step process, not by a translation between radically different forms.

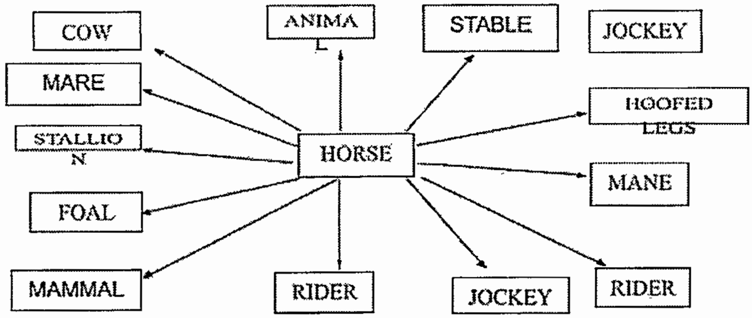
d) Finally, common structures facilitate language learning and conceptual creativity. In learning, a child generalizes conceptual structures learned from experience to form the initial lexical structures needed for language. Metaphor and conceptual refinement create new conceptual structures by adapting old lexical structures to novel situations.

The distinction between lexical structures and conceptual structures provides a principled basis for partitioning knowledge into the lexicon and the more detailed knowledge about the world. Conceptual graphs provide formalism for representing both kinds of structures with a level of precision that allows deeper and more systematic analysis of the relationship between them. As a result, they can help to replace vague discussion with a precise methodology that has a greater chance of being computerized. Finally, the direct mapping between conceptual graph and natural language can simplify the task of knowledge acquisition: a knowledge base of conceptual graphs could be generated directly from natural language inputs. After being primed with a dictionary of lexical knowledge, the system could build up its own encyclopaedia of the world with the aid of a tutor communicating in English, not a knowledge engineer coding in a specialized notation.

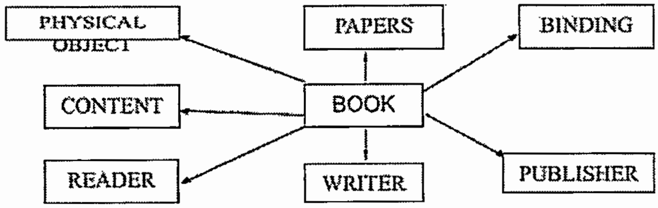
## 11. Expanding the lexical knowledge

The representations we have shown earlier are simple. But we may need more knowledge about lexical items if we are truly making use of them for communication and proper understanding. Take for example, the lexical items *horse* and *book*. They may need at least the following representations. In this representation only concepts related to the main concepts are given in the box. The relations connecting the main concepts with the related concepts are not given. They can be very well understood.

HORSE




BOOK




## Conclusion

The above detailed conceptual representations of lexical knowledge of concepts (realized by words of lexical items in a language) can be mapped into any language. There could be an automatic way of converting these graphs into descriptions in the concerned languages.

## REFERENCE

Cecato, Silvio. (1961). Linguistic Analysis and Programming for Mechanical Translation. New York: Gordon and Beach.

Peirce, Chrales Sander. (1897-1906). Manuscripts on existential graphs. vol. 4, 320-410. Paris

Pustejovksy, J. (1995). The Generative Lexicon. Cambridge, Massachusetts: The MIT Press.

Pustejovsky, J. and B. Boguraev. (1993). "Lexical Knowledge Representation and Natural Language Processing," Artificial Intelligence 63:193-223. Paris

Sowa, J.F. (1984). Conceptual Structures: Information Processing in Mind and Machine. Reading, etc.: Addison-Wesley Publishing Company. Toronto,USA

Vossen P. (eds.) (1999) "EuroWordNet: a multilingual database with lexical semantic networks for European Languages". Dordrecht: Kluwer Academic Publishers.

Vossen P. (1999). "EuroWordNet as a multilingual database". In: Wolfgang Teubert (ed). Berlin: Mouton Gruyte