An Universal Ontology of Unique Atomic and Semantically Compount Concepts

Hermann Bense
bense.com GmbH, Dortmund, GERMANY, hb@bense.com

According to the Wikipedia definition a concept is a ***unit of knowledge created by a unique combination of characteristics***. Concepts are used as the representation of an abstract or concrete thing of the real world in the mind of a human. In ontologies the naming and use of concepts often is cumbersome due to the different meanings words can have, e.g. ***class, bank, ball*** etc. My approach combines methods of semiotics and computer linguistics to overcome this shortfalls. The focus of the work is to build an universal ontology of semantically unique concepts, which are either atomic/base concepts or semantically compound concepts. As an absolutey minimal set of base concepts I took those 60 concepts of [NSM] into account. Also it was promising, that the Basic English of Ogden used only 900 words and that the Longman Defining Vocabulary [Long2014] consists of only 2,000 words. The list of approximately 6,500 Bliss symbols and Bliss words [Bliss1949, Bliss1965] were very helpful in finding out, which symbols/concepts were used to define other symbols/concepts and how often a symbol/concept was used in definition of a Bliss word [BlissUsage]. The atomic concepts for the universal ontology of concepts (UOC) have been carefully selected from the previous list of words so that the derived concepts are disambiguated. Finally, from a set of about 450 atomic concepts all other concepts can be composed.

The construction patterns for concepts uses reification [RDF-Reification,RDF-Vocabulary] to compose binary trees of concepts from sub concepts [John1983]. E.g. a ***manometer*** is a ***gauge*** for ***air pressure***, while ***air_pressure*** is composed from ***air*** and ***pressure.*** The information is internally represented with OWL/RDFS-triples, e.g. *(manometer, <>Subject, gauge)* and *(manometer, <>Object, air_pressure)*, Each atomic/basic concept is associated with an integer number. Applying Cantor´s bijective and total pairing function $\pi(y,x) = y + (x + y)(x + y + 1) / 2$ [CPF, PF], for each binary tree of a concept an integer can be computed, which uniquely encodes all sub concepts. The representation of the concepts is also language independent. For the resulting concept numbering system (CNS), a number of decisions had to be made. This did not only affect the freedom of abritarily assignment of integers to base concepts. But also decisions had to be made, which concept to derive from others, since one would get the same result, if one would chose *north = opposite of south* or *south = opposite of north*. During the computation of concept numbers, conflicts can appear in that way, that a concept number has already been assigned to a basic/atomic concept. In this case the number of the basic/atomic concept will be increased by one and the computation of all concept numbers is repeated. This step is iterated until all conflicts have been resolved.

Encoding concepts as binary trees of sub concepts enbables a another benefit a novel Search by Meaning (SbM) method, which allows to search concepts by any combination

of sub concepts contained therein. Another benefit is, that the sub concepts can be directly used to tag documents.

References

- [Bliss1949] Charles K. Bliss, ***Semantography, a non-alphabetical symbol writing, readable in all languages; a practical tool for general international communication, especially in science, industry, commerce, traffic, etc., and for semantical education, based on the principles of ideographic writing and chemical symbolism***, Sydney: Institute for Semantography. OCoLC: 26684585, **1949**
- [Bliss1965] Charles K. Bliss, ***Semantography (Blissymbolics). 2d enlarged edition. A simple system of 100 logical pictorial symbols, which can be operated and read like 1+2=3 in all languages***, Sydney: Semantography (Blissymbolics) Publications. OCoLC: 1014476., **1965**
- [John1983] Philip N. Johnson-Laird, ***Mental Models: Toward a Cognitive Science of Language, Inference and Consciousness***, Harvard University Press., **1983**, ISBN: 0521273919
- [Long2014] Chris Fox, Rosalind Combley, ***LONGMAN Dictionary of Contemporary English***, Pearson Education Limited, **2014** WebLinks

Weblinks

- [BlissChain2018] *Graph of all BlissWords containing a special BlissSymbol* **http://predicator.name/p1/Ontopage/BlissChain/mind.html**, last visited: 05/02/2018
- [BlissSymbolics2004] *The document describes the Blisssymbolics language, and outlines the rules necessary to be followed for creating new vocabulary* http://eridanus.cz/ id32402/ jazyk/ ume%282le%281_jazyky/bliss_symboly/bliss-rules.pdf, last visited: 05/02/2018
- [BlissUsage] *Frequency of Usage of BlissSymbols in Bliss Words* http://predicator.name/p1 /BlissSymbolics/BlissSymbols/BlissUsages/index.html, last visited: 05/02/2018
- [CPF] *Cantorsche Paarungsfunktion* **https://de.wikipedia.org/wiki/Cantorsche_Paarungsfunktion**, last visited: 05/02/2018
- [NSM] *Natural Semantic Metalanguage* https://en.wikipedia.org /wiki/Natural_ semantic_ metalanguage, last visited: 01/06/2017
- [PF] *Pairing function* **https://en.wikipedia.org/wiki/Pairing_function**, last visited: 01/07/2017
- [RDF-Vocabulary] *The Resource Description Framework (RDF) is a general-purpose language for representing information in the Web*. **http://www.w3.org/TR/2004/REC-rdf-schema-20040210/**, last visited: 05/02/2018
- [RDF-Reification] *RDF Reification allows to describe RDF statements.* http://www.w3.org/TR /2004/REC-rdf-primer-20040210/#reification, last visited: 05/02/2018