Chapter X

How to Represent the World: Ontology-Controlled Natural Languages

Introduction

As far as human knowledge about the world is commonly given in NL expressions and as far as universal ontology is a general science of the world, the examination of its impact on natural language science and technology is among the central topics of many academic workshops and conferences. Ontologists, knowledge engineers, lexicographers, lexical semanticists, and computer scientists are attempting to integrate top-level entity classes with language knowledge presented in extensive corpora and electronic lexical resources. Such a deep quest is mostly motivated by high application potential of reality-driven models of language for knowledge communication and management, information retrieval and extraction, information exchange in software and dialogue systems, all with an ultimate view to transform the World Wide Web into a machine-readable global language resource of world knowledge, the Onto-Semantic Web.

One of the practical applications of integrative ontological framework is to discover the underlying mechanisms of representing and processing language content and meaning by cognitive agents, human and artificial. Specifically, to provide the formalized algorithms or rules, whereby machines could derive or attach significance (or signification) from coded signals, both natural signs obtained by sensors and linguistic symbols.

Of all the valuable outputs, the most significant application of ontological knowledge seems to be computer processing of natural languages, as man-made systems of linguistic forms and categories expressing thoughts and representing reality in all its complexity. The reason why language engineering has been failing to produce intelligent computer systems understanding and generating human language is because of the underestimation of the role

Copyright © 2008, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

of fundamental ontology in the study of natural language (theoretical linguistics), and its main parts: grammar, syntax, and semantics.

One of the misconceptions widely spread both in the KR communities and NLP interest groups is that a human's knowing of the world is basically performed without natural language, and that linguistic communication only marginally involves thinking, "language is froth on the surface of thought," as J. McCarthy states. This view resounds another bad confusion: natural language cannot be knowledge and reasoning representation language, but only an interface, for it is non-algorithmic, ambiguous, vague, and unsystematic. However, such a serious misunderstanding has been met with stiff opposition: "natural language is the ultimate knowledge representation language," (Sowa, 2000a) and that "the richest source of ontological categories is the vocabulary of natural languages" (Sowa, 1995).

Besides, interesting works have appeared promoting NL as a knowledge representation and reasoning system, attributing to it both representational and inferential features and so introducing a new class of large-scale, general-purpose knowledge systems based on natural language (Iwanska & Shapiro, 2000). This NL paradigm of intelligent systems has made the special tracks of international conferences devoted to the theme "NL based Knowledge Representation" (NL, 2005, 2006). Endorsing the language role in human and machine acquisition and processing of information about the world, we support this natural position by devising a formal unified theory of language as a single body of formal rules of grammar and semantics and ontology.

Within the general semantic framework provided by the ontological entities, the formal description of natural language, its elements and structure, is to be based on a consistent set of linguistic expressions and rules:

- Morphologic formulas for the formation of words as linguistic signs and symbol strings expressing content and real meanings
- Phrasal rules for the formation of word strings denoting concrete individuals
- Syntactic rules (or sentential functions) for forming linguistic strings mapping associations and relationships of things
- Linguistic rules of inference forming sentence sequences according to a real logic or ordering of entities

Suggesting a general theory of signs and languages within the ontological paradigm, we establish the following: (1) human language in its nature is tractable algorithmically and systematically; (2) it is expressive and precise; (3) NL is a natural (material) edition and variation of thought; (4) NL is a variant representation of the world knowledge. As much as ontological primitives are knowledge standards, NL symbols and signs are primitives for various knowledge constructions. When considered as a general knowledge and representation language, an opportunity for various NL-driven meaningful machines opens up: NL reasoning applications, NL representation systems, NL answering systems, NL Web search engines, and many other linguistic intelligent programs.

Any powerful knowledge system or intelligent application will be employing a dynamic global model of reality, serving as a common code of real meanings and fundamental axioms,

Copyright © 2008, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

42 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/represent-world-ontology-controlled-natural/28317

Related Content

Review on Latest Trending Topic Detection in Twitter With Stream Processing (Using Fission Pattern)
Saili Ashok Gavhane, Shubham Babanrao Bhadaveand Vengatesan K. (2019). *International Journal of Applied Evolutionary Computation (pp. 43-47).*
www.irma-international.org/article/review-on-latest-trending-topic-detection-in-twitter-with-stream-processing-using-fission-pattern/225042

Buffer Management in Cellular IP Networks using Evolutionary Algorithms
Mohammad Anbarand Deo P. Vidyarthi (2012). *Principal Concepts in Applied Evolutionary Computation: Emerging Trends (pp. 254-275).*
www.irma-international.org/chapter/buffer-management-cellular-networks-using/66824

Virtual Machine Placement Using Statistical Mechanism in Cloud Computing Environment
Narander Kumarand Surendra Kumar (2018). *International Journal of Applied Evolutionary Computation (pp. 23-31).*
www.irma-international.org/article/virtual-machine-placement-using-statistical-mechanism-in-cloud-computing-environment/209415

Ant Colony Programming: Application of Ant Colony System to Function Approximation
Mariusz Boryczka (2010). *Intelligent Systems for Automated Learning and Adaptation: Emerging Trends and Applications (pp. 248-272).*
www.irma-international.org/chapter/ant-colony-programming/38458

Statistical Methods and Artificial Neural Networks Techniques in Electromyography
Ahmad Taher Azarand Valentina E. Balas (2012). *International Journal of System Dynamics Applications (pp. 39-47).*
www.irma-international.org/article/statistical-methods-artificial-neural-networks/62869