A Cognitive Framework for Core Language Understanding and its Computational Implementation

L. Massey, Royal Military College, Canada

ABSTRACT

The author argues that the cognitive processes underlying language understanding may not be logico-deductive or inductive, at least not for basic forms of understanding such as the ability to determine the topics of a text document. To demonstrate this point, they present a human cognition inspired framework for core language understanding and its computational implementation. The framework exploits word related knowledge stored in Long Term Memory (LTM) as well as Short Term Memory (STM) limited capacity, neuromorphic spreading activation and neural activation decay to derive the topics of text. The computational model implementing the framework shows the potential of the approach by establishing that the topics generated by the model are as good as those generated by humans.

Keywords: Cognitive Process, Long Term Memory (LTM), Natural Language Understanding, Short Term Memory (STM), Text Mining, Topics Identification

1. INTRODUCTION

Human language understanding is generally characterized as a logico-deductive process with an underlying inductive learning procedure. For instance, syntactic and semantic analyses, among others, rely heavily on deductive logical structures to infer meaning. As well, the observation of utterances by children is believed to configure the neural language apparatus inductively. However, learning theories cannot explain how children can learn language so well and so rapidly given the inherently small amount of data available to them. This lack of learning data has resulted in the proposal that there may be innate language structures in the brain (Chomsky, 1969, 2006).

Despite this known deficiency of learning data, with the large quantities of electronic texts and increasing computational power, computational linguists have in the last two decades sought meaning in the enormous amount of co-occurrence information found in corpora. Applications of this idea can be found, for instances, in corpus-based statistical methods like Latent Dirichlet Allocation (LDA) (Blei et al., 2003), Latent Semantic Analysis (LSA) (Landauer & Dumais,
Although interesting from a technological application point of view, these techniques are unlikely to be representative of the human cognitive processes of language understanding. Indeed, it is implausible that humans achieve understanding by relying on the presence of large quantities of documents in some sort of mental corpora, or learn by observing and generalizing on thousands of previously classified exemplars. One might claim that reproduction of natural processes or biological plausibility is not a requirement for Artificial Intelligence (AI) or Natural Language Understanding (NLU). On the contrary, one might argue that biological and human cognitive processes might be a valuable source of inspiration for research. It is the position taken in this paper. In that regard, instead of investigating artificial intelligence, we rather study natural intelligence and the underlying cognitive processes involved in language understanding.

There are also real and important practical reasons to motivate our alternative approach. The first one is that existing techniques, by relying on large datasets, are tightly coupled with this data. Coupling is generally recognized as an undesirable situation in software engineering, because it causes modules inter-dependence and greatly increases the complexity of maintenance (Pressman, 2004). Similarly in machine learning and artificial intelligence, dependence on data makes intelligent systems less robust and adaptive when facing changing situations, since they are built according to a static specification and specific view of the world represented by the data (Massey, 2005, 2008). However, human language and conceptualizations are constantly changing (Steels, 2006). It therefore appears futile to statistically induce concepts and meaning from large language datasets (Steels, 2007). The second (and related) problem is autonomy (the ability to function without human intervention). Existing systems, by depending on learning from large collections of language data, are not autonomous since this data must be collected, organized and otherwise prepared by humans. Moreover, following deployment, such systems must be maintained – that is, re-trained on new datasets - when the environment changes. Supervised learning systems are particularly susceptible to this situation, but unsupervised methods are not immune either since they require all sorts of human interventions, for example parameter settings and labeling of extracted clusters. It doesn’t seem to make much sense that so-called intelligent systems would be lacking both autonomy and adaptability, two qualities most people would surely expect to observe from anything qualified of ‘intelligent.’

To solve these problems, we introduce a Cognitive Informatics (CI) (Feldman & Sanger, 2006) framework for basic NLU inspired by how the human mind might process information when one reads. CI aims at investigating intelligence in a transdisciplinary way (Wang, 2003), drawing from computer science, cognitive science, and brain science. CI is strongly influenced and inspired by brain information processing mechanisms and processes, and as such the study of natural intelligence is central to this discipline. Furthermore, CI is interested by practical engineering applications to support society, something we share given our interest in building more effective knowledge management systems. In this paper we do not make a full exploitation of the mathematical constructs introduced by CI, such as Denotational Mathematics (DM), mathematical structures dealing with high-level entities such as abstract objects and complex relations (Wang, 2008) and many others (Wang, 2011). We rather focus on introducing and describing algorithmically a simple fundamental cognitive mechanism for NLU, leaving formalization for future work. We point out that the existing conceptualization of language understanding, namely its logical foundations, may not be required by natural systems to process linguistic information. Nevertheless, modeling them rigorously and formally using for instance the mathematical tools of CI would be scientifically advantageous. Moreover, as it relates to NLU, Wang (2010) identifies a possible synergy between computational linguistics and CI. Wang raises interesting
Duality of Natural and Technological Explanations
www.igi-global.com/chapter/duality-of-natural-and-technological-explanations/127478?camid=4v1a