

# Conversational text composition through commonsense knowledge

Angel Rendon<sup>[0000–0003–3900–9582]</sup>

Universidad Nacional de Colombia  
amrendonsa@unal.edu.co  
<http://unal.edu.co/>

**Abstract.** Natural Language Processing techniques allows us to process text in wide range ways, making possible to extract key information out of texts and even proposing machine translators systems. One of those possibilities is tied to having well trained systems to have smart enough conversations with humans. This work aims to analyze the state-of-the-art techniques and implement them in the construction of a system that using different methods, make it possible to sustain a basic conversation on general topics.

**Keywords:** commonsense knowledge · natural language processing · machine learning · semantic association

## 1 Introduction

One of the artificial intelligence (AI) keystones would be definitively be having fully conversational systems to interact with people for several applications ranging from recommender systems, expert systems, to specialized chatbots and assistants <sup>1</sup>.

Several techniques based on Machine Learning (e.g. Bayesian models, SVM, supervised and unsupervised learning methods), and statistical model methods (e.g. word frequency, text rank, and inverse document frequency), have been used for a long time, with promising results.

However, systems based on these techniques rely on well formed corpora. As an example, WordNet [7] has the following synset for *cat*:

S: (n) computerized tomography, computed tomography, CT, computerized axial tomography, computed axial tomography, CAT (a method of examining body organs by scanning them with X rays and using a computer to construct a series of cross-sectional scans along a single axis)

This simple example could lead to think that it would be possible to miss that specific synset when talking about *computerized tomography* when using

---

<sup>1</sup> <https://www.youtube.com/watch?v=d40jgFZ5hXk>

*cat* in a medical text, showing instead the most probable definition *feline*. That scenario is plausible since the knowledge is strongly dependent on the quality of either unstructured texts or its scale and domain-specific knowledge [4].

Lexical semantic understanding, sustained by socially shared commonsense knowledge on the core as it has the content of what people intends to know while conversing [13].

This contribution aims to:

1. integrates existing work on the semantic association for construction of commonsense knowledge.
2. build an English conversational system using semantic association based on commonsense knowledge construction techniques.
3. assess the performance of the built system compared to traditional conversational approaches.

## 2 Related work

A key element to enrich understanding of sentences is tied to word senses. This fact is more important when the word belongs to a specific domain context. Through hybrid clustering techniques [8] chooses the best terms elements to build the initial committees (i.e. clusters of meanings for a word) out of the WordNet corpus [7], and removing the sense so through further iterations the algorithm can effectively analyze other synsets.

State-of-the-art pre-trained models capture commonsense knowledge with limited value for domain-specific contexts. [3] proposes a general model for capturing specific domain knowledge of software engineering through a word2vec model exploiting available information on Stack Overflow posts. Their work shows that the model effectively disambiguates metaphorical use of English words when it comes to this specific domain, capturing well grained relations within it.

Distributional Semantic Models are another technique for supporting semantic understanding for natural language processing. These models are highly dependent on the size and quality of the corpora that has the commonsense knowledge for the comprehensive task. While English do have high quality and large scaled commonsense and domain specific information, other languages lack enough material to build comprehensive distributional models. [1] proposes to combine lightweight machine translation model using the English Distributional Semantic Model for building enriched knowledge word vectors for other languages. By building a model leveraged by a unigram-level source-target probabilities which can be directly computed from the parallel corpora, the author gets word vectors for other languages from English DSM as feasible activity, showing up that about 66% of improvement has been gotten from lightweight models compared to other models. Spanish got the best performance with nearly 60%, whereas Dutch got the worst with nearly 50%. This demonstrates that lightweight machine translation models is, in the worst case, equivalent (in some

cases outperforming) to the state-of-the-art machine translation systems for translation of word pairs.

As stated at the beginning, one of the important milestones on AI is having intelligent systems to perform common activities as an agent spawned at any location performing a set of actions to answer a question. These actions might require the agent to navigate, process images, understand natural language and learn. [2] proposes a new AI task, EmbodiedQA<sup>2</sup>, where an agent is dropped anywhere and must navigate processing images to gather as much information to answer questions about the environment. To achieve this, the space problem is scoped out using reinforcement learning to reward good trajectories and navigation on well instrumented training environments to provide computer vision, semantic annotations. Their work shows that the navigation problem can be split into two components: The planning (i.e. selecting actions), and the controller (i.e. executing primitive actions); as the *hierarchical navigation module*. Their contribution led to the introduction of the EQA dataset for visual questions of navigation using House3D.

The natural language understanding for stories, as a different kind of narrative, is challenging because it is comprised of a series of events. [12] presents a keyword extraction weight based on word-level attention on external knowledge base. External knowledge must be digitalized to guide the encoding process. Then text keyword extraction is done by combining statistically-based Term Frequency and Inverse Document Frequency, and machine-learning-based K-means. Finally an analysis of sentiment is done by a binary tag (positive and negative) using the VADER classifier to estimate the polarity of a sentence in the binary tag numerical value.

This work aims to work on the lexical semantic understanding, but to fully understand what it does imply, awareness about the cognitive issues in natural languages processing must be present. On their work, [9], address how the cognitive science and the natural language processing fields can help the other by bringing advances on their studies.

“Knowledge can be seen as the result of the confrontation of our a priori ideas with the reality of the outside world” [9]. This leads to some problems:

1. The task-space is probably infinite, since humans can perceive likewise infinite points of view about things.
2. Complex inferences on commonsense knowledge result from perception interference on previous gathered information.
3. The knowledge about how the human brain processes the information is, though promising, very short, and it hardens the proper formalization.

To ease the work commonsense knowledge for natural language processing, a basic assumption must be done about the knowledge: it must be isolated from perception. This assumption led to the creation of the general domain ontology library ConceptNet [11]. However, these kind of databases lack accuracy formalizing specific domain contexts. Nevertheless, these very same databases

<sup>2</sup> <https://embodiedqa.org/>

have led to amazing discoveries ‘just’ by finding patterns revealed by machine learning means [9].

On the other hand, both, cognitive science and natural language processing, have evolved independently for more than thirty years, in part because the technology was not properly advanced, in part because there was not enough information as we do currently have, in part because both fields were not well related [9].

Lately, some research works have shown some convergence of models coming from both study fields like ACT-R theory<sup>3</sup> where the brain work is thought in a modular fashion and how those modules interact to comprehend. In the context of language it does mean a sentence can be analyzed by several modules getting scores of the meaning of words and sentences and how putting all that information together brings more formalized and understood information [9].

Other works focus on the support of two big ideas: **Memory** as the ‘ability to store information and bring it to the front whenever’ it’s needed [9], and **Expectation**, as ‘the predetermination of a lexical category depending on the context’ [9]. This might be treated by research works like [6], stating that long short-term memory (LSTM) are convenient for systems to learn implicit knowledge upon data. Other works exploit this idea in order to measure complexity through large-scale probabilistic context-free grammars that consider both dimensions of the problem [5].

Other work has focused on the understanding of how children acquire language knowledge by hierarchical Bayesian models, or impact on language abilities produced by clinical conditions as Alzheimer’s disease [9].

Graph theory has been used to model lexicon as an intricate network where words and concepts are seen as nodes interconnected to others by specific relations (e.g. proximity in sentences, synonyms, hypernyms hyponyms, synsets, etc) [9].

Although not consider for this work, awareness must be raised for the very adaptive nature of languages. Languages evolve to use new coined words, giving new meanings to existing ones, adopting terms from other dialects or languages [9].

Semantic understanding leveraged by commonsense knowledge is essential to natural language processing tasks. Recent neural models proposed conversation generation based on exclusively conversational data, without deep understanding of users’ input. The training used for these models is either unstructured text or domain-specific knowledge threefold tuples (i.e. entities). As stated before, this makes the model to be tightly-coupled to the data-scale and quality. The contribution of [13], presents a commonsense knowledge aware conversational model that retrieves some knowledge graphs out of the knowledge base to enrich the text meaning and respond informatively and appropriately. Using both, an static and a dynamic graphs attention mechanisms a better understanding of the text is achieved. The former is used to enrich the semantic representation of the text. The latest reads not only the knowledge graphs but the entities (as

---

<sup>3</sup> Adaptive Control of Thought-Rational

threefold tuples, or triples for short) using the semantic information from graphs and entities to generate better responses. The following are the main results of this work:

1. Using a large-scale commonsense knowledge in neural conversation generation to understand and generate more appropriate responses improves the responses quality.
2. Using entities not like isolated elements, but as a whole through a graph leads to a better interpretation of the text semantics.

To investigate the impact of commonsense knowledge in the context of semantic association between lexical concepts, [10] proposes a work done by extending and integrating preliminary works on different aspects such as the perception of similarity with lexical entities and disambiguation of ConceptNet [11] instances. Then three main research questions were formulated to validate:

1. if functional/behavioral information improves perception of meaning and semantic association.
2. if it was possible to integrate commonsense knowledge into the WordNet lexicon and determine if such integration improves accuracy of semantic association classifiers.
3. if using commonsense knowledge helps label and explain semantic association.

## References

1. Barzegar, S., Davis, B., Handschuh, S., Freitas, A.: Multilingual Semantic Relatedness Using Lightweight Machine Translation. Proceedings - 12th IEEE International Conference on Semantic Computing, ICSC 2018 **2018-Janua**, 108–114 (2018). <https://doi.org/10.1109/ICSC.2018.00024>
2. Das, A., Datta, S., Gkioxari, G., Lee, S., Parikh, D., Batra, D.: Embodied Question Answering. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 1–10. IEEE (jun 2018). <https://doi.org/10.1109/CVPR.2018.00008>, <https://ieeexplore.ieee.org/document/8578106/>
3. Efstathiou, V., Chatzilenas, C., Spinellis, D.: Word embeddings for the software engineering domain. In: Proceedings of the 15th International Conference on Mining Software Repositories - MSR '18. pp. 38–41. ACM Press, New York, New York, USA (2018). <https://doi.org/10.1145/3196398.3196448>, <http://dl.acm.org/citation.cfm?doid=3196398.3196448>
4. Ghazvininejad, M., Brockett, C., Chang, M.W., Dolan, B., Gao, J., Yih, W.t., Galley, M.: A Knowledge-Grounded Neural Conversation Model pp. 5110–5117 (2017), <http://arxiv.org/abs/1702.01932>
5. Hale, J.: A probabilistic earley parser as a psycholinguistic model (2007). <https://doi.org/10.3115/1073336.1073357>
6. Ma, Y., Peng, H., Khan, T., Cambria, E., Hussain, A.: Sentic LSTM: a Hybrid Network for Targeted Aspect-Based Sentiment Analysis. Cognitive Computation **10**(4), 639–650 (2018). <https://doi.org/10.1007/s12559-018-9549-x>

7. Miller, G.A., Miller, G.A., Beckwith, R., Fellbaum, C., Gross, D., Miller, K.: WordNet: An on-line lexical database. *INTERNATIONAL JOURNAL OF LEXICOGRAPHY* (1990)
8. Pantel, P., Lin, D.: Discovering word senses from text **41**, 613 (2004). <https://doi.org/10.1145/775047.775138>
9. Poibeau, T., Vasishth, S.: Cognitive Issues in Natural Language Processing **55**(2), 7–15 (2016). <https://doi.org/10.1201/b21583-2>, <http://arxiv.org/abs/1610.07365><http://dx.doi.org/10.1201/b21583-2>
10. Ruggeri, A., Di Caro, L., Boella, G.: The Role of Common-Sense Knowledge in Assessing Semantic Association. *Journal on Data Semantics* **8**(1), 39–56 (2019). <https://doi.org/10.1007/s13740-018-0094-2>, <https://doi.org/10.1007/s13740-018-0094-2>
11. Speer, R., Chin, J., Havasi, C.: ConceptNet 5.5: An Open Multilingual Graph of General Knowledge (Singh 2002) (2016), <http://arxiv.org/abs/1612.03975>
12. Yuan, Q., Tong, J., Gao, S., Li, S.: Story Understanding with External Knowledge Based Attention. *Proceedings of 2018 6th IEEE International Conference on Network Infrastructure and Digital Content, IC-NIDC 2018* pp. 229–233 (2018). <https://doi.org/10.1109/ICNIDC.2018.8525570>
13. Zhou, H., Young, T., Huang, M., Zhao, H., Xu, J., Zhu, X.: Commonsense knowledge aware conversation generation with graph attention. *IJCAI International Joint Conference on Artificial Intelligence* **2018-July**, 4623–4629 (2018)