Hierarchical Ontology Graph for Solving Semantic Issues in Decision Support Systems

Hua Guo and Kecheng Liu
Informatics Research Centre, University of Reading, Reading, U.K.

Keywords: Ontology Graph, NLP, Knowledge Graph, Decision Support Systems, Semantic Composition, Neural-symbolic Integration.

Abstract: In the context of the development of AI algorithms in natural language processing, tremendous progress has been made in knowledge abstraction and semantic reasoning. However, for answering the questions with complex logic, AI system is still in an early stage. Hierarchical ontology graph is proposed to establish analysis threads for the complex question in order to facilitate AI system to further support in business decision making. The study of selecting the appropriate corpora is intended to improve the data asset management of enterprises.

1 INTRODUCTION

Strategic questions are often formulated as openly as possible in order to stimulate more considerations involving different perspectives of business operations. Those questions are not easy to answer and usually demand a great deal of effort of analysis before they can be addressed adequately. In order to achieve the goal to help managers identify the hidden impact factors of decision making, the enormous academic explorations over query understanding (Moldovan et al., 1999), information retrieval and process (Harman,1993) or over heterogeneous data sources (Kumar et al., 2014) have been ongoing for years. Along with the solutions of key issues for knowledge engineering, e.g.: semantic parsing (Cai and Yates, 2013), knowledge graph (Kwiatkowski et al., 2013), have been proposed in recent years, the open-domain question answering system is becoming one of the most important applications in AI-NLP arenas. Most state-of-the-art searches adopts the bottom-up approach, which firstly classifies the question into a few classes (Moldovan et al.,1999; Li and Roth, 2002), secondly retrieve the relative information (Stoyanchev et al, 2008), and finally extract the answer from the relative documents (Ravichandran and Hovy, 2002). The purpose of the question processing stage is to narrow down the search scope and it focuses on question classification (Allam and Haggag, 2012). For the open-domain QA system, which relies on universal ontology and information, the bottom-up approach handles well. However, the closed-domain questions, like the strategic questions related to the business operation, usually consist of multiple aspects and contain complex judgment logic. It is often the case that two of the keywords in completely unrelated fields are logically linked when answering a particular question. This really is a challenge for AI algorithms to abstract the tacit relationship from the limited corpus. So, how can these AI techniques help in answering strategic questions? In this paper, a top-down approach of hierarchical ontology graph is proposed, which is starting from the question analysis. This approach will embed the logic of business operation and the collaborative relationships of departments in the organization into the procedure of decomposing user queries into sub-questions. In other words, this approach focuses on question reformulation (Allam and Haggag, 2012) to understand the query in an enterprise context and transform the complex logic question into a few simple logical questions to empower the search engine. From a practical point of view, the hierarchical ontology graph is a visualized procedure of decision-making, and the threads of analysing the strategic questions could be mapped to the entities on the hierarchical ontology graph.

Hierarchical ontology graph, rooted in the knowledge graph (Singhal, 2012), illustrates complex question from the relevant aspects, breaks one general question into several domain-level sub-questions, and

Guo, H. and Liu, K.
Hierarchical Ontology Graph for Solving Semantic Issues in Decision Support Systems.
DOI: 10.5220/0007769904830487
In *Proceedings of the 21st International Conference on Enterprise Information Systems (ICEIS 2019)*, pages 483-487
ISBN: 978-989-758-372-8

gradually decomposes the domain-level questions layer by layer until the sub-questions can be answered by the existing databases/documents. The reasoning procedure will display on a hierarchical ontology graph.

2 STATEMENT OF POSITION

A thorough understanding of the company operational details is a prerequisite for decision making. However, those details dispersed among organizations and isolated within domains. The routine approach of an executive making decision is to call a meeting involving the heads of all departments (domain experts), who have the ability to interpret or decompose the strategic question into sub-questions in domain-level, and their subordinates can decompose these sub-questions into queries and bridge these queries with the existing database or documents. With the help of frontline staff or data analysts, these queries will be answered, and bringing together the answers to these queries, the subordinates can answer sub-questions to their department head. As these sub-questions are answered individually and then aggregated, finally the strategic question could be answered. This is the normal procedure of an executive makes a strategic decision, which is not only inefficient but also significantly impacted by the personal experience of domain experts, reflected in the ability to interpret the question. Even more, interest disputes are also likely to arise between departments due to the knowledge barriers.

Hierarchical ontology graph can speed up this process and can bridge the knowledge islands. It can break down complex problem layers into simple questions that can be answered by existing data sources, and it can also provide an enterprise-wide knowledge graph that effectively eliminates knowledge barriers between departments. A shared and reusable knowledge graph is an effective tool to construct an agile organization in the quickly evolving business competitive environment. It will help on rapidly engaging in multidirectional communication and complex collaboration.

3 PROPOSED SOLUTION

The proposed hierarchical ontology graph consists of four different levels, shown as in Figure 1. The application ontology is constructed by the data of department-level, which theoretically docks with the data warehouse, indicating the analysis results and reports of the current business performance. Task ontology is a department-level knowledge graph abstracted from the internal corpora, which refer to the documents of enterprise processes and operational logics. Considering the number of relevant internal documents is comparatively limited, it is highly recommended to annotate these documents as much as possible in order to apply the full supervision or semi supervision learning algorithms. Different to task ontology, domain ontology is built on the industry professional corpora, which may or may not be well trained. In order to get the enterprise-level knowledge graph, the outcomes got from training the semantically-rich annotated corpora in department level can be used to do semi supervision training. The semi-automatic tools, like ONION (Mitra et al., 2000), are also recommended to bridge department-level ontologies in the procedure of creating domain ontologies. The top layer in the figure is a top-level ontology, which includes some tacit knowledge. These inexplicit factors could be abstracted from the business activities based on the strategic management theory and may vary in industries.

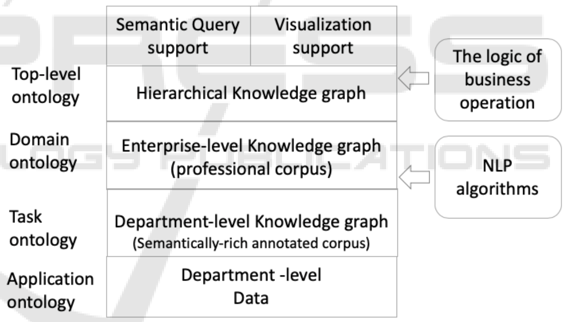

Figure 1: Hierarchical ontology graph.

Considering that the number of the corpus in a specific domain is usually very limited in practice, deep learning algorithms, which is based on a large-scale corpus, often find the difficulty to obtain high-quality training results. The latest research direction of neural-symbolic integration network (Garcez et al., 2008) is trying to combine the strengths of connective and symbolic paradigms to enhance the ability of machine learning and reasoning. The framework of Object-oriented Neural Programming (OONP) (Lu et al., 2017) is proposed for semantically parsing documents in specific domains, which leverage the advantages of reasoning feature of the symbolic network to construct object-oriented ontology in the process of text comprehension. The OONP framework provides another approach to constructing the hierarchical ontology graph, furthermore, the design of carry-on memory (Lu et al., 2017) model can effectively store and reuse the prior knowledge, which can be used in dealing with the existing business logic in enterprise ontology graph. Evans and Grefenstette (2018) proposed a logic programming method based on Inductive Logical Programming (ILP) to reason on symbolic domains, which could effectively support the reasoning function of a hierarchical ontology graph.

The following example will explain how a hierarchical ontology graph support decision making on a specific strategic question.

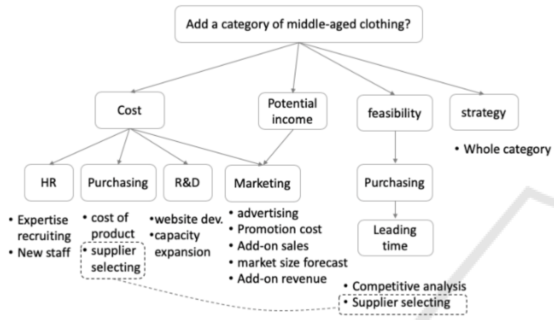

Figure 2: Decomposing a complex question over hierarchical ontology graph.

Suppose the CEO of an e-commerce company is considering adding a new category of middle-aged clothing. His question is “Whether should we add a category of middle-aged clothing?”. Cost and income might be the first two key factors to be taken into consideration because all the business activities revolve around the purpose of making a profit. So, primarily this business-related question type can be decoded into a ‘quantitative’ question in terms of cost and potential income. Furthermore, profitability is not the only indicator of business decision-making, but also the need to take into account the company's long-term strategic plan. Feasibility is another dimension to evaluate the possibility of success of the project, and it is listed as a separating factor in the top-level ontology as well.

At the second level of the ontology graph, the four aspects sub-questions have been allocated to the different domains. They are HR department, R&D department, purchasing department, and marketing department. In practice, entities at this level are often closely related to the organizational structure.

At the third level of the ontology graph, the same indicator may appear in different domains, such as “supplier selecting” appears in both “cost” and “feasibility”. The same indicator can map to multiple domains and also one concept can be interpreted into different meanings. In this case, “cost” has different implications in HR, purchasing, R&D and marketing departments. There are 1-to-N and N-to-1 tree-like relationships and even can emerge N-to-N network relationships in some cases. At this point, at the third level, a complex strategic issue has been broken down into a variety of queries belonging to various departments.

The fourth level of the ontology graph does not show in Figure 2, which are the answers to the queries in the third level and got from the data analysis of the existing database.

To build an enterprise-level knowledge graph, selecting appropriate corpora is one of the most important tasks besides selecting algorithms or models. Corpus selection will be one of the future research directions of this research.

The above description presents a decomposition process of how hierarchical ontology graph performs on complex questions, and AI algorithms will help on constructing the hierarchical ontology graph.

Artificial intelligence is changing the business landscape, especially with the development of weak supervision learning and self-learning neural in NLP. Usually, the result is highly correlated with the quality of the corpus. Comparing with the general linguistic corpus, the corpus in a business context has clear boundaries and are comparatively simple in terms of less ambiguity, polysemy, and vagueness issues. The workload of getting a semantically-rich annotated corpora is manageable, which is a crucial impact factor of the computing result.

For the corpus with rich semantic annotations, the full supervision training models can be applied, e.g.: DeepCoder (Balog, 2016), NPI (Reed & Freitas, 2015) and Seq2Tree (Dong & Lapata, 2016). To train the small volume of databases, the end-to-end models like Neural Programmer (Neelakantan et al, 2015) and Neural Turing Machines (Graves et al, 2014) can be adopted. For those context-sensitive processing, LSTM (Hochreiter & Schmidhuber 1997), Seq2Seq (Sutskever et al, 2014), named entity recognition (Lample et al, 2016), and reading Comprehension models (Yu et al., 2018) can be considered. Those NLP related algorithms provide the feasible methods for concepts extraction from text and from semi-structured tables (Pasupat & Liang 2015).

The complex question defined in this paper referrers to a question that contains multiple independent variables, rather than the complex syntax logic of the sentence. These independent variables are often difficult to obtain from ready-made documents, and they exist in the minds of domain experts in the form of knowledge or experience. The human knowledge composed in different forms, including tree-like relationships, two-dimensional grid relationships, single dimensional sequential relationships and directed grid relationships (Kemp and Tenenbaum, 2009). The classical TransE (Bordes et al., 2013) model and its derivations are not strong enough to present these cognitive models from a mathematical approach. This is the reason why a symbolic network will be considered to do relationship reasoning and neural networks will focus on learning and information extraction. The hierarchical ontology graph proposes an approach to building the close-domain question and answering system by leveraging the prior experiences to support decision making.

4 SUGGESTED COURSES OF ACTION

The hierarchical ontology graph is proposed to solve semantic issues through injecting business operation logic and the experiences of domain experts to support executives to make strategic decisions. The procedure of constructing an enterprise-level ontology graph is also the process of establishing the organizational knowledge graph. A unified knowledge graph can not only help on decision making but also be the basis for efficient business operation. Further research will include the following aspects:

1. Enterprise Semantic Model: constructing the abductive reasoning model for decision support
2. Algorithm: selecting appropriate algorithms to match the requirements for semantic analysis
3. Corpus acquirements: working out which types of documents in an enterprise can be trained as corpus
4. Tacit knowledge transfer: visualizing the tacit enterprise experience in a hierarchical ontology graph.

REFERENCES

Allam, A.M.N. and Haggag, M.H., 2012. The question answering systems: A survey. International Journal of Research and Reviews in Information Sciences (IJRRIS), 2(3).

Balog, M., Gaunt, A.L., Brockschmidt, M., Nowozin, S. and Tarlow, D., 2016. Deepcoder: Learning to write programs. arXiv preprint arXiv:1611.01989.

Bordes, A., Usunier, N., Garcia-Duran, A., Weston, J. and Yakhnenko, O., 2013. Translating embeddings for modeling multi-relational data. In Advances in neural information processing systems (pp. 2787-2795).

Cai, Q. and Yates, A., 2013. Large-scale semantic parsing via schema matching and lexicon extension. In Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) (Vol. 1, pp. 423-433).

D. K. Harman. Overview of the first Text REtrieval Conference (TREC-1). In Proceedings of the First Text REtrieval Conference (TREC-1), pages 1–20. NIST Special Publication 500-207, March 1993.

Dong, L. and Lapata, M., 2016. Language to logical form with neural attention. arXiv preprint arXiv:1601.01280.

Evans, R. and Grefenstette, E., 2018. Learning explanatory rules from noisy data. Journal of Artificial Intelligence Research, 61, pp.1-64

Garcez, A.S.A., Lamb, L.C. and Gabbay, D.M., 2008. Neural-symbolic cognitive reasoning. Springer Science & Business Media.

Graves, A., Wayne, G. and Danihelka, I., 2014. Neural turing machines. arXiv preprint arXiv:1410.5401.

Hochreiter, S. and Schmidhuber, J., 1997. Long short-term memory. Neural computation, 9(8), pp.1735-1780.

Kemp, C. and Tenenbaum, J.B., 2009. Structured statistical models of inductive reasoning. Psychological review, 116(1), p.20.

Kwiatkowski, T., Choi, E., Artzi, Y. and Zettlemoyer, L., 2013. Scaling semantic parsers with on-the-fly ontology matching. In Proceedings of the 2013 conference on empirical methods in natural language processing (pp. 1545-1556).

Lample, G., Ballesteros, M., Subramanian, S., Kawakami, K. and Dyer, C., 2016. Neural architectures for named entity recognition. arXiv preprint arXiv:1603.01360.

Li, X. and Roth, D., 2002, August. Learning question classifiers. In Proceedings of the 19th international conference on Computational Linguistics-Volume 1 (pp. 1-7). Association for Computational Linguistics.

Liu Zhiyuan, Sun Maosong, Lin Yankai, et al., 2016. Knowledge Representation Learning: A Review[J]. Journal of Computer Research and Development, 53(2): 247-261.

Lu, Z., Liu, X., Cui, H., Yan, Y. and Zheng, D., 2017. Object-oriented neural programming (oonp) for document understanding. arXiv preprint arXiv:1709.08853.

Mitra, P., Wiederhold, G. and Kersten, M., 2000, March. A graph-oriented model for articulation of ontology interdependencies. In International Conference on Extending Database Technology (pp. 86-100). Springer, Berlin, Heidelberg.

Moldovan, D., Harabagiu, S., Pasca, M., Harabagiu, A., Mihalcea, R., Girju, R., Goodrum, R. and Rus, V., 1999. Lasso: A tool for surfing the answer net.

Neelakantan, A., Le, Q.V. and Sutskever, I., 2015. Neural programmer: Inducing latent programs with gradient descent. arXiv preprint arXiv:1511.04834.

Panupong Pasupat and Percy Liang. 2015. Compositional semantic parsing on semi-structured tables. In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), pages 1470–1480.

Ravichandran, D. and Hovy, E., 2002, July. Learning surface text patterns for a question answering system. In Proceedings of the 40th annual meeting on association for computational linguistics (pp. 41-47). Association for Computational Linguistics.

Reed, S. and De Freitas, N., 2015. Neural programmer-interpreters. arXiv preprint arXiv:1511.06279.

Singhal, A ‘‘Introducing the knowledge graph: Things, not strings,’’ May 2012. [Online]. Available: https://www.blog.google/products/search/introducing-knowledge-graph-things-not/.

Suresh kumar, G., Zayaraz, G., 2014. Concept relation extraction using Naıve Bayes classifier for ontology-based question answering systems. J. King Saud Univ.

Sutskever, I., Vinyals, O., and Le, Q. V. (2014). Sequence to sequence learning with neural networks. Technical report. arXiv:1409.3215 [cs.CL] Google. NIPS’2014.

Stoyanchev, S., Song, Y.C. and Lahti, W., 2008, August. Exact phrases in information retrieval for question answering. In Coling 2008: Proceedings of the 2nd workshop on Information Retrieval for Question Answering (pp. 9-16). Association for Computational Linguistics.

Yih, S.W.T., Chang, M.W., He, X. and Gao, J., 2015. Semantic parsing via staged query graph generation: Question answering with knowledge base.

Yu, Adams, Wei, David Dohan, Quoc Le, Thang Luong, Rui Zhao, and Kai Chen. 2018. Fast and accurate reading comprehension by combining self-attention and convolution. In International Conference on Learning Representations.