Improving the Accuracy of Black-Box Language Models with Ontologies: A Preliminary Roadmap

Marco Monti[1,2,*], Oliver Kutz[1], Guendalina Righetti[3] and Nicolas Troquard[4]

[1]Free University of Bozen-Bolzano (UNIBZ), Piazza Università, 1—39100, Bozen-Bolzano (BZ), Italy
[2]IBM, Circonvallazione Idroscalo, 20090, Segrate (MI), Italy
[3]University of Oslo, Department of Philosophy, Oslo, Norway
[4]Gran Sasso Science Institute (GSSI), Viale Francesco Crispi, 7—67100, LAquila (AQ), Italy

Abstract

Large Language Models (LLMs) have revolutionised natural language generation. But their statistical and auto-regressive nature makes them unreliable. It has become clear to the research community that in order to produce reliably correct answers, LLMs need to be enriched in some way with 'world models' reflecting the semantics of the domains being queried. We here propose a simple workflow to address this problem through a neuro-symbolic interaction protocol with the LLM treated as a blackbox. Answers given by an LLM are checked against accepted knowledge provided by a domain ontology. The approach aims to combine conflict detection with explanation extraction and formal repairs presented to the LLM in the form of specific artificial speech acts. The goal is to build constraining, incremental prompts that improve repeatability and veracity in the LLM's output.

Keywords

LLM, ontologies, neuro-symbolic reasoning

1. Introduction

LLMs offer serious potential in knowledge discovery and information retrieval given the training on vast corpora of text. Examples include simple lookups of basic facts, summarisation in the style of Wikipedia abstracts, producing reformulations of difficult-to-understand documents such as those found in medical diagnosis, etc. [1]. It is therefore no surprise that some of the apparent skills of LLMs are also being explored for complementing or assisting ontological reasoning tasks, namely in particular learning new subsumptions and building concept taxonomies, or populating existing ontologies with entities [2, 3].

Limitations. Despite their remarkable achievements, LLMs also exhibit a number of significant drawbacks. Like other models of artificial neural networks, LLMs are susceptible to

Proceedings of the Joint Ontology Workshops (JOWO) - Episode X: The Tukker Zomer of Ontology, and satellite events co-located with the 14th International Conference on Formal Ontology in Information Systems (FOIS 2024), July 15-19, 2024, Enschede, The Netherlands.

*Corresponding author.
Marco.Monti@student.unibz.it (M. Monti); oliver.kutz@unibz.it (O. Kutz); guendalina.righetti@ifikk.uio.no (G. Righetti); nicolas.troquard@gssi.it (N. Troquard)
0009-0001-3308-4752 (M. Monti); 0000-0003-1517-7354 (O. Kutz); 0000-0002-4027-5434 (G. Righetti); 0000-0002-5763-6080 (N. Troquard)

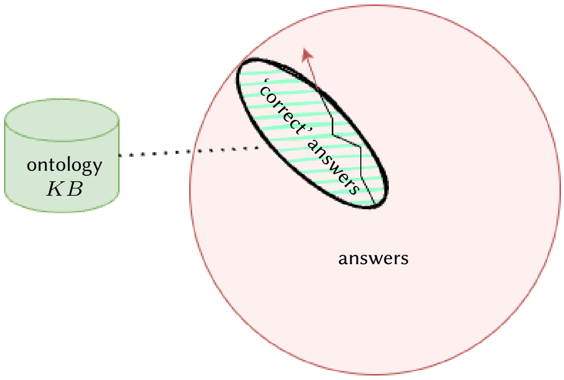

Figure 1: LLMs answers tend to exit the space of correct answers. A domain ontology and reasoning can help detecting it when it occurs.

biases and have limited contextual understanding [4]. But perhaps the most critical limitation concerns their lack of accuracy, manifested by so-called hallucinations [5, 6].

Yann LeCun's "unpopular opinion" in a recent series of talks has been that "auto-regressive LLMs are doomed" [7]. An argument for this claim goes roughly as follows. Assume a language model M, and suppose that p is the probability that any token produced by M takes us outside of the set of correct answers. The probability that an answer of length n provided by the language model M is correct is then $(1-p)^n$, which converges to zero with the length of the answer. Of course, with a sufficiently small value of p (close to 0) and a 'not too long' answer, the performance of such a model M can still be very high. Figure 1 illustrates this general situation. Our core concern thus is to study how to use ontologies and formal reasoning to steer the LLM to remain in, or at least close to, the space of 'correct' answers. We first provide some context on existing mitigation methods for hallucinations which are knowledge-based.

Some existing hallucination mitigation methods. Yin et al. [8] propose a generative question answering system. To provide correct answers, the system is connected to a triple store of true facts, from which it retrieves a set of candidate facts and generates an answer to the question. Further methods exploring how to enhance machine reading comprehension systems by incorporating external knowledge sources are presented by Bi et al. [9]. Li et al. [10] also address the issue of semantic drift in generative question answering by incorporating external knowledge. Martino et al. [11] use knowledge injection to counter hallucinations in large language models. Retrieval-Augmented Generation (RAG) [12] uses expert knowledge and domain-specific related documents to augment answers to queries, which are then processed together by the LLM to better contextualize them.

Ji et al. [13] propose a method for mitigating LLM hallucinations via self-reflection. This approach involves three self-reflective loops: factual knowledge acquisition, knowledge-consistent answering, and question-entailment answering. Galitsky [14] presents a fact-checking system that exploits web mining to find correct information to suggest to the LLM. The system capitalises on argumentation analysis and defeasible logic programming to handle inconsistent sources. For a more complete survey of existing hallucination mitigation methods, refer to [15].

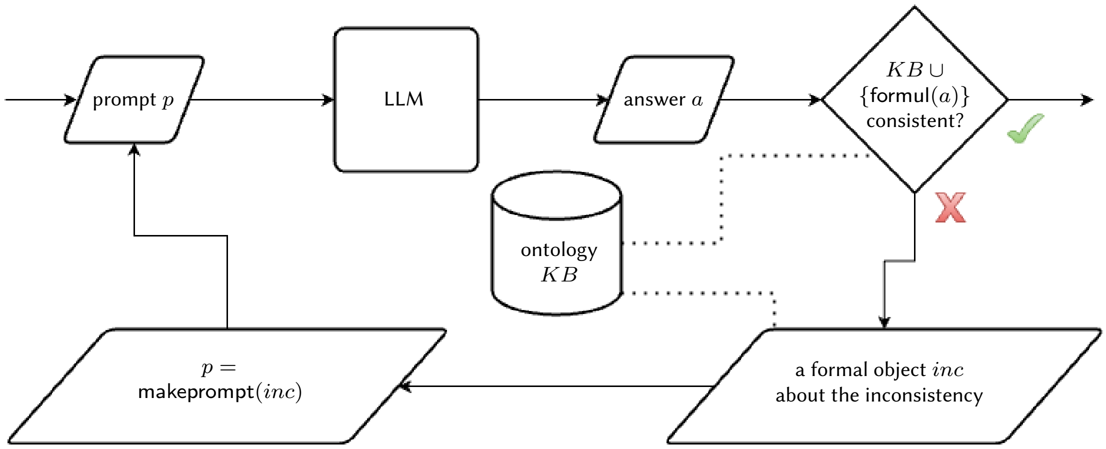

Figure 2: A generic workflow to steer language models on the path of accurate answers.

Using formal ontologies as world models. It has become clear that LLMs need world models to generate factual responses. Ontologies are natural candidates for being the providers of these world models.

Ontologies are formal descriptions of the entities present within a domain of interest. They describe the concepts, the individuals that populate them, and the relationships that hold between them. Ontologies may contain implicit knowledge (a set of general axioms) and explicit knowledge (a set of factual statements). The latter can be represented as rows in a relational database, as a knowledge graph, or a Description Logic (DL) ABox. The former must be represented as a set of logical formulas, e.g., as a DL TBox. Description Logics are natural candidates because their reasoning problems (consistency checking, entailment, instance checking, etc.) are usually decidable and efficient algorithms and implementations exist. Yet, they maintain a reasonable expressivity. DLs also form the theoretical underpinning of the W3C Web Ontology Language (OWL). See [16] for an introduction to Description Logics. Moreover, some of the key technical elements that are required to interact and converse with an LLM are more readily available in the DL context, as described further below. This includes non-classical reasoning approaches for conflict detection, knowledge debugging, formal argumentation, or knowledge weakening [17].

We thus suggest to use logical reasoning with background ontological domain knowledge to detect inconsistent answers (Figure 1), and iteratively nudge the LLMs' answers back on a path of correct answers.

2. Steering LMs towards accuracy: a generic workflow

A circular workflow is depicted in Figure 2, illustrating the possible enhancement of a Language Model's capabilities through interfacing with ontological reasoning. This workflow iteratively enriches the LLM's inference abilities by fostering a symbiotic relationship between linguistic proficiency and ontological reasoning.

We briefly discuss the key elements of our workflow:

Prompt p**:** The interaction begins with the user providing input in the form of text. This input can in general range from simple queries to complex prompts, questions, or commands. In our scenario, the prompt may be designed to:

1. ask for a succinct answer, so as to limit the issue with auto-regressive LLMs,
2. use only simple concepts and relations that have a direct counterpart in the given ontology,
3. target certain central concepts in the specific domain of knowledge according to subject-matter-experts.

LLM: The LLM is treated as a blackbox. The workflow does not interfere with learning and is not intended for fine-tuning purposes.

Answer a**:** a is a textual response to the prompt p generated by the LLM. This response a typically appears to be a coherent and relevant piece of text that addresses the user's prompt. Ideally, this answer is short, as requested by p, to limit the issue with auto-regressive LLMs.

Formulizer formul**:** is a computational module designed to convert English responses a into formal expressions $\mathsf{formul}(a)$ represented (largely) within the signature and logical language used in the KB. This problem has been addressed by the DL research community [18, 19]. Transforming a textual response into a given formal language can of course be also achieved through appropriately training a network [20].

Coherence validation module: This module is aimed at evaluating aspects of consistency and coherence of the answer a to prompt p. Besides the answer a, it also takes as input a domain ontology KB related to the topic of the prompt, here expected to be written in DL and using certain known concepts, roles, and individuals.

We assume here that the outcome a is in a form that makes it amenable to semantic analysis in order to generate a formal response, as follows:

1. *LLM outcome:* The generated text by the LLM, a, follows a structure and vocabulary that allows one to extract a formalised version $\mathsf{formul}(a)$ written in the same language as KB.
2. *Semantic analysis:* $KB \cup \{\mathsf{formul}(a)\}$ is evaluated for semantic defects which include inconsistency but also weaker notions such as 'off topic' or 'incoherent'.
3. *Coherence evaluation:* The module may provide a coherence score (or other quantitative evaluation metrics) indicating the degree of alignment or agreement between the LLM outcome, domain ontology, and other constraints. Such metrics should helps assess the quality and reliability of the generated text and help steer the feedback. Scores for coherence were for instance proposed in [21].
4. *Feedback inc about incoherence:* If incoherences are detected, the module may provide feedback inc highlighting the areas where the LLM outcome diverges from the ontology or violates logical rules. This feedback can be used to refine the generated text or improve the LLM's understanding of the domain.

Verbalizer verbal**:** The use of verbalisation techniques for translating symbolic facts, ontological rules and logic entailments into natural language is a core aspect of the workflow. Verbalisations are readily available for the DL framework; in the simplest form, Manchester Syntax can be almost directly translated to regular English [22].

Prompt adaptation $p = \mathsf{makeprompt}(inc)$**:** We create a speech act generated from the semantic analysis. This can be a verbalisation of an explanation of a proof, an announcement that certain facts need to be accepted, or that other facts need to be rejected.

Extracting formal explanations is arguably a challenge of its own [23, 24]. Fortunately, we are not interested in a *post hoc* explanation of the response of the LLM, but in logical derivations of the found inconsistency. Some simple forms of 'explanation' can be considered, like the extraction of a minimal inconsistent set, for which there are efficient methods (e.g., [25]). Thus, in the workflow, inc could be for example a minimal inconsistent set of $KB \cup \{\mathsf{formul}(a)\}$, and the new prompt $p = \mathsf{makeprompt}(inc)$ could be "But $\mathsf{verbal}(\neg inc)$!". To further help the LLM, we might want to suggest a repair of the inconsistency, and perhaps some weakened assumptions of the claims that the LLM had made, using, e.g., the repair and weakening techniques of [17].

3. Possible limitations

Our approach itself certainly has some limitations. Some of those possible limitations concern assumptions we are making about the future (capabilities) of LLMs themselves.

- Our workflow involves some sort of (automated) "arguing" with the LLM. We started this note by reporting on the lack of accuracy of LLMs. And yet, we must rely on some accuracy, or at least logical consistency. Indeed, for our workflow to work as expected, the LLM would need to have a basic 'understanding' of logic. (E.g., the updated prompt attempting to point out a contradiction to the LLM: "But, $\mathsf{verbal}(\neg inc)$!") Unfortunately, current LLMs are deficient in this regard [26, 27] and perform especially poorly in the presence of negations [28]. The logical understanding they have is limited to statistical patterns in language rather than true logical comprehension. However, improving exactly this skill is a core research problem in the field [29].
- Abstraction (and, thus, logic) is still very difficult to handle by LLMs, as is clear also from studying their mathematics capabilities [27] but also when describing verbally the description of an object using variables. To illustrate this, Fig 3a shows the result of a text-to-image generation where the text prompt is a verbalisation of a formalisation of the concept 'fishvehicle' as created by a symbolic blending algorithm and using phrases such as 'the object x should be such that [...]'. In contrast, Fig 3b shows the direct text-to-image production using 'a fish that is also a vehicle', relying directly on the bias of the model what a 'fish' respectively a 'vehicle' look like. Both artefacts were produced with *SDXL-Lightning*[1].

[1]See https://huggingface.co/spaces/ByteDance/SDXL-Lightning.

(a) A symbolic representation of a 'fishvehicle' produced by the blending algorithm of [33], verbalised and fed back into a text-to-image generation algorithm.

(b) A textual representation of a 'fishvehicle', namely described as 'a fish that is also a vehicle', fed directly into a text-to-image generation algorithm.

Figure 3: Representations of 'fishvehicle'.

- LeCun [7] recommends to abandon auto-regressive LLMs for text generation altogether. Our approach does not act upon this recommendation. Instead it tries (maybe naively) to put back text generation on the path of truth.

To summarise, to be put in practice, one needs the following elements:

- A domain ontology in a signature S_D for every domain D that is addressed in prompt p.
- A reasoner for the specific Description Logic (or corresponding OWL profile) in which the domain ontology is written. Fortunately, consistency checking is a standard reasoning task, and efficient reasoners exist for DLs (e.g., Hermit, Fact++, Pellet), but also to some extent for First Order Logic (e.g., Vampire, Z3).
- A verbalizer verbal, to transform a set of logical formulas in Description Logic over the signature S_D into natural language. Some readily available technologies have been proposed for *controlled natural languages*. Examples are the OWL-Verbalizer [30] for Attempto Controlled English (ACE),[2] or the mapping proposed by Cregan et al. [31] between the Sydney OWL Syntax and OWL 1.1 functional syntax. The NaturalOWL System [32] is specifically aimed for generating coherent multi-sentence translations of OWL axioms.
- A formulizer formul, to transform English text of the domain D in a logical representation, in Description Logic, over the signature S_D. If we can assume that the answer provided by the LLM is in controlled English, then the verbalizer, like Kaljurand's OWL-Verbalizer is reversible, meaning that it can convert ACE English back into (ACE) OWL.

4. Outlook

We are interested in addressing the challenge of making LLMs more reliable. This paper lays the groundwork for future research by proposing a preliminary roadmap. To this end, we have proposed a high-level architecture for interacting with an LLM through a conversational pipeline

[2]See also https://www.w3.org/2001/sw/wiki/ACE.

that incorporates artificial speech acts, including feedback from symbolic components. To fully assess the potential of this architecture, concrete examples and instantiations are needed.

Acknowledgments

The authors thank Ruslan Idelfonso Magaa Vsevolodovna for valuable feedback.

References

[1] S. Minaee, T. Mikolov, N. Nikzad, M. Chenaghlu, R. Socher, X. Amatriain, J. Gao, Large language models: A survey, arXiv preprint arXiv:2402.06196 (2024).

[2] M. Funk, S. Hosemann, J. C. Jung, C. Lutz, Towards ontology construction with language models, arXiv preprint arXiv:2309.09898 (2023).

[3] H. Babaei Giglou, J. D'Souza, S. Auer, LLMs4OL: Large language models for ontology learning, in: International Semantic Web Conference, Springer, 2023, pp. 408–427.

[4] I. O. Gallegos, R. A. Rossi, J. Barrow, M. M. Tanjim, S. Kim, F. Dernoncourt, T. Yu, R. Zhang, N. K. Ahmed, Bias and fairness in large language models: A survey, CoRR abs/2309.00770 (2023). URL: https://doi.org/10.48550/arXiv.2309.00770. doi:10.48550/ARXIV.2309.00770. arXiv:2309.00770.

[5] S. Lin, J. Hilton, O. Evans, TruthfulQA: Measuring how models mimic human falsehoods, in: S. Muresan, P. Nakov, A. Villavicencio (Eds.), Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Association for Computational Linguistics, Dublin, Ireland, 2022, pp. 3214–3252. URL: https://aclanthology.org/2022.acl-long.229. doi:10.18653/v1/2022.acl-long.229.

[6] H. Alkaissi, S. McFarlane, Artificial Hallucinations in ChatGPT: Implications in Scientific Writing, Cureus 15 (2023).

[7] Y. LeCun, Objective-Driven AI: Towards Machines that Can Learn, Reason, and Plan, 2024. AAAI/IAAI Invited Talk.

[8] J. Yin, X. Jiang, Z. Lu, L. Shang, H. Li, X. Li, Neural generative question answering, in: M. Iyyer, H. He, J. Boyd-Graber, H. Daumé III (Eds.), Proceedings of the Workshop on Human-Computer Question Answering, Association for Computational Linguistics, San Diego, California, 2016, pp. 36–42. URL: https://aclanthology.org/W16-0106. doi:10.18653/v1/W16-0106.

[9] B. Bi, C. Wu, M. Yan, W. Wang, J. Xia, C. Li, Incorporating external knowledge into machine reading for generative question answering, in: K. Inui, J. Jiang, V. Ng, X. Wan (Eds.), Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), Association for Computational Linguistics, Hong Kong, China, 2019, pp. 2521–2530. URL: https://aclanthology.org/D19-1255. doi:10.18653/v1/D19-1255.

[10] C. Li, B. Bi, M. Yan, W. Wang, S. Huang, Addressing semantic drift in generative question answering with auxiliary extraction, in: C. Zong, F. Xia, W. Li, R. Navigli (Eds.), Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 2: Short

Papers), Association for Computational Linguistics, Online, 2021, pp. 942–947. URL: https://aclanthology.org/2021.acl-short.118. doi:10.18653/v1/2021.acl-short.118.
[11] A. Martino, M. Iannelli, C. Truong, Knowledge injection to counter large language model (LLM) hallucination, in: C. Pesquita, H. Skaf-Molli, V. Efthymiou, S. Kirrane, A. Ngonga, D. Collarana, R. Cerqueira, M. Alam, C. Trojahn, S. Hertling (Eds.), The Semantic Web: ESWC 2023 Satellite Events - Hersonissos, Crete, Greece, May 28 - June 1, 2023, Proceedings, volume 13998 of *Lecture Notes in Computer Science*, Springer, 2023, pp. 182–185. URL: https://doi.org/10.1007/978-3-031-43458-7_34. doi:10.1007/978-3-031-43458-7_34.
[12] P. S. H. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W. Yih, T. Rocktäschel, S. Riedel, D. Kiela, Retrieval-augmented generation for knowledge-intensive NLP tasks, in: H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, H. Lin (Eds.), Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual, 2020. URL: https://proceedings.neurips.cc/paper/2020/hash/6b493230205f780e1bc26945df7481e5-Abstract.html.
[13] Z. Ji, T. Yu, Y. Xu, N. Lee, E. Ishii, P. Fung, Towards mitigating LLM hallucination via self reflection, in: H. Bouamor, J. Pino, K. Bali (Eds.), Findings of the Association for Computational Linguistics: EMNLP 2023, Association for Computational Linguistics, Singapore, 2023, pp. 1827–1843. URL: https://aclanthology.org/2023.findings-emnlp.123. doi:10.18653/v1/2023.findings-emnlp.123.
[14] B. A. Galitsky, Truth-O-Meter: Collaborating with LLM in Fighting its Hallucinations, Preprints (2023). URL: https://doi.org/10.20944/preprints202307.1723.v1. doi:10.20944/preprints202307.1723.v1.
[15] Z. Ji, N. Lee, R. Frieske, T. Yu, D. Su, Y. Xu, E. Ishii, Y. J. Bang, A. Madotto, P. Fung, Survey of hallucination in natural language generation, ACM Comput. Surv. 55 (2023). URL: https://doi.org/10.1145/3571730. doi:10.1145/3571730.
[16] S. Rudolph, Foundations of Description Logics, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011, pp. 76–136. URL: https://doi.org/10.1007/978-3-642-23032-5_2. doi:10.1007/978-3-642-23032-5_2.
[17] N. Troquard, R. Confalonieri, P. Galliani, R. Penaloza, D. Porello, O. Kutz, Repairing Ontologies via Axiom Weakening, in: Proc. of the Thirty-Second AAAI Conference on Artificial Intelligence, AAAI Press, 2018.
[18] R. R. de Azevedo, F. Freitas, R. Rocha, J. A. A. de Menezes, L. F. A. Pereira, Generating description logic $\mathcal{ALC}$ from text in natural language, in: T. Andreasen, H. Christiansen, J.-C. Cubero, Z. W. Raś (Eds.), Foundations of Intelligent Systems, Springer International Publishing, Cham, 2014, pp. 305–314.
[19] B. Gyawali, A. Shimorina, C. Gardent, S. Cruz-Lara, M. Mahfoudh, Mapping natural language to description logic, in: The Semantic Web: 14th International Conference, ESWC 2017, Portorož, Slovenia, May 28–June 1, 2017, Proceedings, Part I 14, Springer, 2017, pp. 273–288.
[20] A. Ranta, Translating between language and logic: What is easy and what is difficult, in: N. Bjørner, V. Sofronie-Stokkermans (Eds.), Automated Deduction – CADE-23, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011, pp. 5–25.

[21] R. Confalonieri, O. Kutz, N. Troquard, P. Galliani, D. Porello, R. Peñaloza, M. Schorlemmer, Coherence, Similarity, and Concept Generalisation, in: A. Artale, B. Glimm, R. Kontchakov (Eds.), Proc. of the 30th International Workshop on Description Logics (DL 2017), volume 1879, CEUR-WS, Montpellier, France, July 18–21, 2017, 2017.
[22] M. Horridge, N. Drummond, J. Goodwin, A. Rector, R. Stevens, H. Wang, The Manchester OWL Syntax, in: OWL: Experiences and Directions, 2006.
[23] W. Saeed, C. W. Omlin, Explainable AI (XAI): A systematic meta-survey of current challenges and future opportunities, Knowl. Based Syst. 263 (2023) 110273. doi:10.1016/J.KNOSYS.2023.110273.
[24] S. Ali, T. Abuhmed, S. H. A. El-Sappagh, K. Muhammad, J. M. Alonso-Moral, R. Confalonieri, R. Guidotti, J. D. Ser, N. D. Rodríguez, F. Herrera, Explainable artificial intelligence (XAI): what we know and what is left to attain trustworthy artificial intelligence, Inf. Fusion 99 (2023) 101805. URL: https://doi.org/10.1016/j.inffus.2023.101805. doi:10.1016/J.INFFUS.2023.101805.
[25] K. M. Shchekotykhin, D. Jannach, T. Schmitz, Mergexplain: Fast computation of multiple conflicts for diagnosis, in: Q. Yang, M. J. Wooldridge (Eds.), Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, IJCAI 2015, Buenos Aires, Argentina, July 25-31, 2015, AAAI Press, 2015, pp. 3221–3228. URL: http://ijcai.org/Abstract/15/454.
[26] M. E. Jang, T. Lukasiewicz, Consistency Analysis of ChatGPT, 2023. arXiv:2303.06273.
[27] S. Frieder, J. Berner, P. Petersen, T. Lukasiewicz, Large language models for mathematicians, 2024. arXiv:2312.04556.
[28] T. H. Truong, T. Baldwin, K. Verspoor, T. Cohn, Language models are not naysayers: an analysis of language models on negation benchmarks, in: A. Palmer, J. Camacho-collados (Eds.), Proceedings of the 12th Joint Conference on Lexical and Computational Semantics (*SEM 2023), Association for Computational Linguistics, Toronto, Canada, 2023, pp. 101–114. URL: https://aclanthology.org/2023.starsem-1.10. doi:10.18653/v1/2023.starsem-1.10.
[29] L. Pan, A. Albalak, X. Wang, W. Wang, Logic-LM: Empowering large language models with symbolic solvers for faithful logical reasoning, in: H. Bouamor, J. Pino, K. Bali (Eds.), Findings of the Association for Computational Linguistics: EMNLP 2023, Association for Computational Linguistics, Singapore, 2023, pp. 3806–3824. URL: https://aclanthology.org/2023.findings-emnlp.248. doi:10.18653/v1/2023.findings-emnlp.248.
[30] K. Kaljurand, Attempto Controlled English as a Semantic Web Language, Ph.D. thesis, Faculty of Mathematics and Computer Science, University of Tartu, 2007. URL: http://attempto.ifi.uzh.ch/site/pubs/papers/phd_kaljurand.pdf.
[31] A. Cregan, R. Schwitter, T. Meyer, Sydney OWL syntax - towards a controlled natural language syntax for OWL 1.1, in: C. Golbreich, A. Kalyanpur, B. Parsia (Eds.), Proceedings of the OWLED 2007 Workshop on OWL: Experiences and Directions, Innsbruck, Austria, June 6-7, 2007, volume 258 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2007.
[32] I. Androutsopoulos, G. Lampouras, D. Galanis, Generating natural language descriptions from OWL ontologies: the naturalowl system, J. Artif. Intell. Res. 48 (2013) 671–715. URL: https://doi.org/10.1613/jair.4017. doi:10.1613/JAIR.4017.

[33] G. Righetti, D. Porello, N. Troquard, O. Kutz, M. Hedblom, P. Galliani, Asymmetric Hybrids: Dialogues for Computational Concept Combination, in: B. Brodaric, F. Neuhaus (Eds.), 12th International Conference on Formal Ontology in Information Systems - FOIS 2021, Frontiers in Artificial Intelligence and Applications, IOS Press, 2021.