Technical Report

1 More Details on NaturalOWL

NaturalOWL [1] is a natural language generation system that produces texts describing individuals or classes of OWL ontologies. Unlike simpler OWL verbalizers which typically express a single axiom at a time in controlled and are not entirely fluent natural language primarily for the benefit of domain experts, NaturalOWL generates fluent and coherent multi-sentence texts for end-users. NaturalOWL adopts a pipeline architecture, which is common in NLG. Its system generates texts in three stages, document planning, micro-planning, and surface realization.

Document planning. First, NaturalOWL produces the message triples to be expressed, with each triple intended to be easily expressible as a single sentence. The text planner of NaturalOWL then orders the message triples, in effect ordering the corresponding sentences. Because most of the sentences simply provide additional information about the target or the second-level targets, NaturalOWL does not consider global coherence.

Micro-planning. Micro-planning consists of three substages: lexicalization, sentence aggregation, and generation of referring expressions. In NaturalOWL, for every property of the ontology and every supported natural language, the domain author may specify one or more template-like sentence plans to indicate how message triples involving that property can be expressed. In our study, we present a set of certain templates for axiomatic-turning sentences. Our experiment of study on alternative translation of axioms demonstrates the robustness of our approach when changing some part of the template.

Surface realization. In NaturalOWL, the ordered sentence plans at the end of micro-planning already completely specify the surface final form of each sentence. The surface realization of NaturalOWL is mostly a process of converting internal, but fully specified and ordered sentence specifications to the final text.

2 Properties of *Similarity*

In this part we prove the three properties of the two proposed similarity calculation methods. Under the given axiom translation rules and certain embedding method, the distributed semantic vector which a certain axiom is transformed into is unique. In the following part, we denote $Emb(ON(\alpha))$ as v_1, $Emb(ON(\beta))$ as v_2.

Proof of the properties of Similarity:
Range: Since $CosineDistance(v_1, v_2) \in [-1, 1]$ then $Similarity_{Cos}(v_1, v_2) \in [0, 1]$. Since $EuclideanDistance(v_1, v_2) \in [0, \infty)$, then $Similarity_{Euc}(v_1, v_2) \in (0, 1]$.Obviously, $Similarity_{Cos}$ is monotonically decreasing with respect to the included angle $\gamma \in [0, \pi)$. $Similarity_{Euc}$ s monotone concerning the Euclid Distance of the two vectors. So the closer the similarity calculation value is to 1, the higher the similarity between the two vectors is. In turn, the closer the similarity calculation value is to 0, the lower the similarity between the two vectors is.
Grammatical Reflexivity: Because of the uniqueness of the transformation, $v_1 = v_2$. The included angle between v_1 and v_2 is 0. So $CosineDistance(v_1, v_2) = 1$ and we conclude that $Similarity_{Cos}(v_1, v_2) = \frac{1}{2}(1+cos(v_1, v_2)) = 1$. Similarly, $Similarity_{Euc}(v_1, v_2) = 0$, then $Sim(\alpha, \beta) = 1/(1 + Similarity(v_1, v_2)) = 1$.
Symmetry: For $Similarity_{Cos}(v_1, v_2) = \frac{1}{2}(1+cos(v_1, v_2)) = \frac{1}{2}(1+cos(v_2, v_1)) = Similarity_{Cos}(v_2, v_1)$ and $Similarity_{Euc}(v_1, v_2) = 1/(1+Similarity_{Euc}(v_1, v_2)) = 1/(1+Similarity_{Euc}(v_1, v_2)) = Similarity_{Euc}(v_2, v_1)$, we can conclude that the two similarity calculation methods satisfy the symmetry.

3 Details on logical properties

We consider the logical nature of the two inference methods defined above. Our consideration is mainly from the minimal set of expected properties of preferential inference relations [24] (also called system P) and the one of rational inference relations [24] (also called system R).

Although we vectorize the entities in all axioms, for a certain maximum consistent subset, its credibility is still the sum of the credibility of all elements of the set. The score of axioms is global and does not change for the change of the set. This is essentially the same as the approach of [23] to scoring axioms.
Proof of Theorem 1.
Necessity. For every consistent set $K_i \subseteq K$, for every nontrivial axiom $\alpha \in K \backslash K_i$, $K_i \cup \{\alpha\} \succ_K K_i$, $\text{score}_{K,\oplus}^{s}(K_i \cup \{\alpha\}) > \text{score}_{K,\oplus}^{s}(K_i)$, $\bigoplus_{\beta \in K_i \cup \{\alpha\}} s(K, \beta) - \bigoplus_{\beta \in K_i} s(K, \beta) = s(K, \alpha) > 0$. It shows that for every non-trivial axiom $\alpha \in K$, there exists a consistent subset $K_i \subseteq K$, such that $\alpha \notin K_i.\alpha$ must be included into at least one $K_n \in \text{mcs}(K)$.
case 1 : The cardinality of $K_n, |K_n| = 1$, then there must be a subset $K_m \in mcs(K), m \neq n$, such that $\alpha \notin K_m$.
case 2: The cardinality of $K_n, |K_n| > 1$, then $\alpha \notin K_n \backslash \{\alpha\}$. Since the arbitrariness of α is proved above, we conclude that for every non-trivial axiom $\alpha \in K, s(K, \alpha) > 0$.
Sufficiency. For every non-trivial axiom $\alpha \in K, s(K, \alpha) > 0$, since $\text{score}_{K,\oplus}^{S}(K_i) = \bigoplus_{\beta \in K_i} s(K, \beta)$, then for each consistent set $K_i \subseteq K, \text{score}_{K,\oplus}^{S}(K_i \cup \{\alpha\}) - \text{score}_{K,\oplus}^{S}(K_i) = \bigoplus_{\beta \in K_i \cup \{\alpha\}} s(K, \beta) - \bigoplus_{\beta \in K_i} s(K, \beta) = s(K, \alpha) > 0$, namely $\text{score}_{K,\oplus}^{S}(K_i \cup \{\alpha\}) > \text{score}_{K,\oplus}^{S}(K_i)$, which means $K_i \cup \{\alpha\} \succ_K K_i$.
Proof of Theorem 2.

Since Theorem 2 is a rewriting of Theorem 5.18 in [24], we refer the reader to the original paper for detailed proof.

Proof of Theorem 3.

Suppose an arbitrary non-trivial axiom $\alpha \in K$. Considering in the Grammatical Reflexivity of *Similarity*:

$$\begin{aligned} s(K,\alpha) &= mc(K,\alpha) \\ &= \sum_{\{K_i \in mcs(K) \mid \alpha \in K_i\}} agg(K_i,\alpha) \\ &= \sum_{\{K_i \in mcs(K) \mid \alpha \in K_i\}} (\frac{1}{|K_i|} \sum_{\beta \in K_i} Sim(\alpha,\beta)) > 0 \end{aligned}$$

So the selection relation based on our proposed method is monotonic relation according to Theorem 1. And Theorem 2 shows the rationality of the corresponding inference relation. Due to the rationality of our proposed method, the reasoning satisfies:

Ref	$\alpha \mid\sim \alpha$	**Cut**	$\dfrac{\alpha \wedge \beta \mid\sim \gamma, \alpha \mid\sim \beta}{\alpha \mid\sim \gamma}$
LLE	$\dfrac{\alpha \leftrightarrow \beta, \alpha \mid\sim \gamma}{\beta \mid\sim \gamma}$	**Or**	$\dfrac{\alpha \mid\sim \gamma, \beta \mid\sim \gamma}{\alpha \vee \beta \mid\sim \gamma}$
RW	$\dfrac{\alpha \rightarrow \beta, \gamma \mid\sim \alpha}{\gamma \mid\sim \beta}$	**CM**	$\dfrac{\alpha \mid\sim \beta, \alpha \mid\sim \gamma}{\alpha \wedge \beta \mid\sim \gamma}$
RM	$\dfrac{\alpha \mid\not\sim \neg\beta, \alpha \mid\sim \beta}{\alpha \wedge \beta \mid\sim \gamma}$		

Ref is Reflexivity.

Cut expresses the fact that one may, in his way towards a plausible conclusion, first add a hypothesis to the facts he knows to be true and prove the plausibility of his conclusion from this enlarged set of facts, and then deduce (plausibly) this added hypothesis from the facts.

LLE is Left Logical Equivalence. Left Logical Equivalence expresses the requirement that logically equivalent formulas have exactly the same consequences.

Or says that any formula that is, separately, a plausible consequence of two different formulas, should also be a plausible consequence of their disjunction.

RW is Right Weakening. Right Weakening obviously implies that one may replace logically equivalent formulas by one another on the right of the $\mid\sim$ symbol.

CM is Cautious Monotonicity. Cautious Monotonicity expresses the fact that learning a new fact, the truth of which could have been plausibly concluded should not invalidate previous conclusions.

RM is Rational Monotonicity. It expresses the fact that only additional information, the negation of which was expected, should force us to withdraw plausible conclusions previously drawn. It is an important tool in minimizing the updating we have to do when learning new information.

Model	Method	AUTO.	biop.	UOBM-35	UOBM-36	UOBM-37	UOBM-38
#mc		0.1	0.6	0.1	0.6	4.5	4.3
SBERT	Cosine	129.2	112.5	28.1	42.5	167.4	164.1 %
	Euclid	121.3	109.5	25.6	41.6	166.0	161.1
RoBERTa	Cosine	127.8	104.8	24.4	39.8	164.5	157.9
	Euclid	125.3	112.5	24.6	37.9	162.4	154.4
ALBERT	Cosine	129.5	112.6	29.6	42.2	174.7	168.6
	Euclid	123.0	110.1	26.9	41.6	172.4	167.3
ConSERT	Cosine	134.9	126.9	32.1	49.3	181.9	177.6
	Euclid	129.2	132.0	31.1	47.6	176.8	174.4
TransE	Cosine	593.6	574.7	452.7	461.1	528.0	512.4
	Euclid	577.8	409.2	449.2	461.5	432.1	434.1
TransH	Cosine	758.9	706.9	613.5	627.5	694.4	697.3
	Euclid	665.4	670.2	610.0	626.3	702.5	703.3
TransD	Cosine	769.8	734.3	625.5	629.5	644.2	710.0
	Euclid	670.6	680.5	623.2	632.6	706.4	710.8
TransR	Cosine	711.9	730.3	616.7	615.5	678.3	682.8
	Euclid	648.4	662.7	613.5	613.4	694.7	684.4
RotatE	Cosine	149.3	172.1	5.7	14.3	85.8	413.8
	Euclid	57.6	39.0	5.9	15.0	86.8	424.3
RDF2Vec	Cosine	368.1	389.2	359.6	361.5	422.4	415.6
	Euclid	350.0	387.2	359.4	361.8	449.8	424.5

Table 1: Study on efficiency (The unit of all data is seconds).

4 Detailed Results on Efficiency

All the evaluation results of efficiency are shown in Table 1. We do not present the results of skeptical inference and CMCS because no selection needs to be executed by Skeptical inference and it is very efficient for CMCS to select those cardinality-maximal subsets (i.e., no more than 1 second for each ontology). For the six ontologies in our experiment, the consumed time of our proposed method is within 20 minutes. Although our proposed methods spend more time than the baselines, they are efficient enough in practice as the selection only needs to be performed once for each ontology and this process can be done offline. For all methods proposed in this paper, they are very efficient and a query can be answered within about half a second.

Overall, for efficiency, Sentence Embedding is better than KG Embedding and Euclid Distance is better than Cosine Distance. Among the models of Sentence Embedding used in our experiment, RoBERTa is the most efficient, and RotatE is the most efficient in KG Embedding methods

5 Detailed Results on Accuracy

Model	Method	UOBM-35	UOBM-36	UOBM-37	UOBM-38
SBERT	Cosine	97.73%	94.95%	96.59%	95.96%
	Euclid	96.59%	96.97%	97.73%	96.97%
RoBERTa	Cosine	98.86%	97.98%	97.73%	94.95%
	Euclid	96.59%	96.97%	97.73%	94.95%
ALBERT	Cosine	98.86%	97.98%	97.73%	94.95%
	Euclid	96.59%	96.97%	97.73%	94.95%
ConSERT	Cosine	98.86%	94.95%	98.86%	94.95%
	Euclid	96.59%	96.97%	97.73%	94.95%
TransE	Cosine	95.45%	94.95%	97.73%	96.97%
	Euclid	95.45%	96.97%	97.73%	94.95%
TransH	Cosine	95.45%	94.95%	98.86%	96.97%
	Euclid	95.45%	96.97%	97.73%	95.96%
TransD	Cosine	95.45%	95.56%	97.73%	94.95%
	Euclid	95.45%	97.98%	97.73%	96.97%
TransR	Cosine	96.59%	95.96%	98.86%	94.95%
	Euclid	95.45%	96.97%	97.73%	95.96%
RotatE	Cosine	96.59%	96.97%	98.86%	96.97%
	Euclid	98.86%	95.96%	100.00%	97.97%
RDF2Vec	Cosine	97.73%	96.97%	94.32%	94.95%
	Euclid	98.86%	95.96%	100.00%	97.97%

Table 2: Evaluation results.

From our previous experimental results, we can see that our method has strong reasoning ability, which is obviously beyond baselines. But besides paying attention to whether a reasoning method can deduce more axioms, we should also pay attention to whether it can accurately deduce correct axioms. Therefore, we selected four inconsistent ontologies UOBM-35, UOBM-36, UOBM-37 and UOBM-38 for experiments. We use the consistent ontology UOBM-lite-10 before inserting the conflict to test the accuracy of reasoning and querying results of the MCS selected by our method. The results are shown in Table 2. The experimental results are all above 94.95%, close to 100%. Although the accuracy of Skeptical Inference and some other methods can reach 100%, their reasoning power is too low. In contrast, our method has strong reasoning ability while maintaining high accuracy.

OWL statements	Origin Sentences	Replacement sentences
SubclassOf(*Target*, *NamedClass*)	*Target* is subclass of *C*	1. Every *Target* is a *C* 2. There is no *Target* that is not a *C*
ClassAssertion(*NamedClass*, *Target*)	*Target* is an instance of *NamedClass*	1. *Target* belongs to *NamedClass* 2. *Target* is part of *NamedClass*
EquivalentClasses(*Target*, *C*)	*Target* is equivalent to *C*	1. *Target* equates to *C* 2. *Target* corresponds to *C*
SubclassOf(*Target* ObjectOneOf(*indiv1* *indev2* ...))	*Target* is one of to *indiv1* *indev2* ...	1. *Target* is an instance of {*indiv1* *indev2* ...}

Table 3: Replacement rules of alternative encoding study.

6 Replacement Rules of Alternative Encoding Study

In Section 6.4 in our paper, We verify the robustness of our method by studying on influence of alternative translation of axioms, which indicates that although the same axiom can be translated into two different sentences, which may violate semantic reflexivity, it has little impact on the performance of our proposed methods. We randomly replace the axiom-to-natural language rule in NatualOWL, and Table 3 shows some replacement rules.