Unpacking Language Categories: Uncertainty and Similarity*

Ares Fabregat-Hernández[1,2], Javier Palanca[1], and Vicent Botti[1,3]

[1] Valencian Research Institute for Artificial Intelligence (VRAIN), Universitat Politècnica de València, Camí de Vera s/n, 46022, Valencia, Spain
[2] Universidad Internacional de Valencia (VIU), C/Pintor Sorolla 21, 46002 Valencia, Spain
[3] valgrAI (Valencian Graduate School and Research Network of Artificial Intelligence)

Abstract. In this short manuscript we propose how to use the semantic categories $\mathcal{L}_T$ and $\mathcal{P}_T$ of a given text T to find the relation of evaluative expressions. Since this kind of expressions heavily rely on the context their are used we introduce a method to understand their meaning by looking at factored endomorphisms of elements in $\mathcal{P}_T$.

1 Introduction

In [1], the authors establish syntax and semantic categories based on a text T to prove some results about logical implications of the meanings of words. More recently, in [2, 3] the authors have introduced the concept of Markov Categories to model uncertainty in transmission processes and divergence of information.

We extend those results to a category $\mathcal{P}_T$ with subsets of expressions as objects and probabilistic matrices as morphisms. With this we can measure the similarity between linguistic expressions in a certain context given by the text. This takes the form of matrices with conditional probabilities representing the semantic closeness of expressions of T. With that, we can define semantic spaces for the expressions and study the logical implications or effects they have on the receiver of the information. This can then be applied to machine learning algorithms to fine-tune their vector representation of the concepts to better their result in machine-human interaction. Furthermore, the inverse is also possible: by understanding the category $\mathcal{P}_T$ we can deduce a great deal about the text T. For instance, we can deduce inherent biases present in the text or the subtle differences in evaluative linguistic expressions for the text T.

* This work is partially supported by the TAILOR project, a project funded by the EU Horizon 2020 research and innovation programme under GA No 952215, and by the GUARDIA project, a project funded by Generalitat Valenciana GVA-CEICE project PROMETEO/2018/002

2 Semantic Categories

Meaning is not an absolute concept: it is relative to the context. This means that in order to represent semantic information of words or expressions we need to fix a context, or a text corpus, denoted by T. Once a text is fixed we can define the enriched category $\mathcal{L}_T$ whose objects are expressions in T and morphisms are arrows $x \to y$ with a probability attached to it. If $x \leq y$, that is, if x is a sub-expression of y then the arrow, denoted $\mathcal{L}_T(x, y) = p(y|x)$ is the conditional probability to find y once we have found x in T. If, on the other hand, x is not a sub-expression of y we define:

$$p(x\|y) = \inf_{g \in \mathcal{L}_T} \left\{ \frac{p(g|x)}{p(g|y)}, 1 \right\}.$$

If x and y are similar, that is, they appear in similar contexts and with similar frequency we can say that they have comparable meanings. This is represented by the probability $p(x\|y)$ which can be interpreted as the likelihood that we could use both expressions interchangeably (using the text T as a reference).

The problem with the category L_T is that the expressions are separated in our representation: each expression g is an object in L_T. This implies that the relations among objects are...

To solve this we aggregate the objects in L_T to form subsets of expressions in T and we define the category $\mathcal{P}_T$ whose objects are subsets of expression in T and morphisms between two sets X and Y are probabilistic matrices. This means that the morphisms are matrices with entries in $[0, 1]$ that represent probabilities. A subcategory of $\mathcal{P}_T$ is the category $\mathcal{M}_T$ with row-stochastic matrices as morphisms which is a Markov Category as in [2, 3].

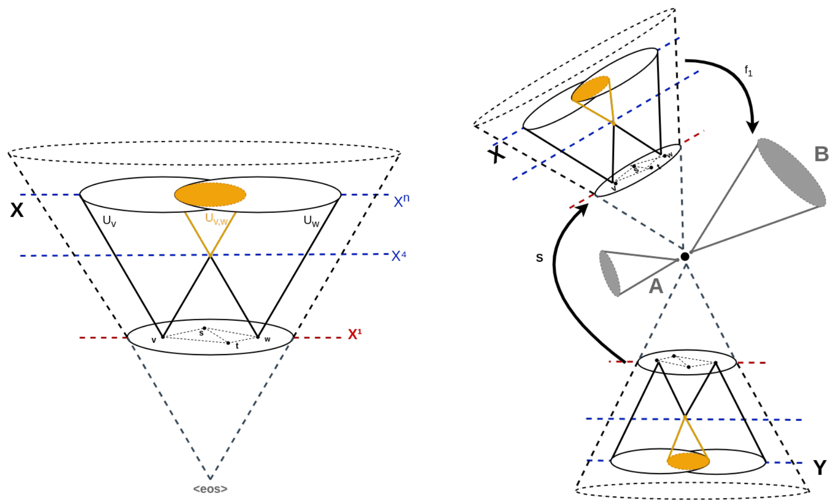
Figure 1 shows a nice way to picture these categories. Sub-figure 1a represents the category $\mathcal{L}_T$ as a cone, starting at an empty symbol <eos> where each expression g in $\mathcal{L}_T$ is the start of the cone whose elements are expressions containing g. Horizontal section of $\mathcal{L}_T$ are called **graded pieces**. For instance the first graded piece contains the expressions of length one in T, that is, the words in T.

Sub-figure 1b represents the category $\mathcal{P}_T$ the set of all possible cones. Morphisms are matrices that change from one cone to another. Thus, matrices are in escence changing the semantic information found in an object in $\mathcal{P}_T$. The cones are called **Alexandrov cones** by their relation the the Alexandrov topological space of a pre-ordered set.

With these categories defined we can begin the study of evaluative linguistic expressions.

3 Logic and Language

Evaluative linguistic expressions are expressions such as small, medium, large, very deep, rather shallow. They represent qualities of subjects that are understandable to human readers but difficult to interpret with machine (mathematical) logic. The interesting part is that they are relative to the context. For



(a) Caption1 (b) Caption1

Fig. 1. The semantic Categories

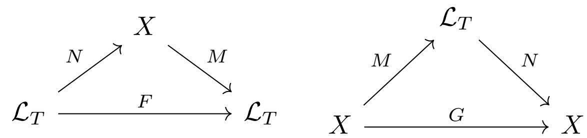
instance, long time or old, are concepts that do not correspond to a specific amount of time. In particle physics, long can be a few seconds whereas in cosmology, a star with millions of years can still be young.

This means that, as before, we must restrict our attention to a specific context, a text T, to derive the meaning of these expressions. Given T and the categories $\mathcal{L}_T$ and $\mathcal{P}_T$, let X be a set of linguistic evaluative expressions, that is an object in $\mathcal{P}_T$. The we can look at the spaces $\mathcal{P}_T(X, \mathcal{L}^1_T)$ of probabilistic matrices from the evaluative expressions to the words of T and $\mathcal{P}_T(\mathcal{L}^1_T, X)$. The composition of morphisms in these spaces is an endomorphism of $\mathcal{L}_T$ which acts like Johari windows of each word.

Specifically, given $M \in \mathcal{P}_T(X, \mathcal{L}^1_T)$ and $N \in \mathcal{P}_T(\mathcal{L}^1_T, X)$ the multiplication $F = MN$ lies in $\mathcal{P}_T(\mathcal{L}^1_T, \mathcal{L}^1_T)$. This is an association of words, or ideas related by their evaluative expressions. Indeed we can interpret F as a family of probability measures F_w over $\mathcal{L}^1_T$ indexed by the elements of $\mathcal{L}^1_T$. Breaking up this composition yields that N can be viewed as a family of probability measures N_w over X indexed by the elements of $\mathcal{L}^1_T$ and M can be viewed as a family of probability measures M_g over $\mathcal{L}^1_T$ indexed by the elements of X. Thus, the entries with higher probability in M correspond to the words we associate more closely with the evaluative expression. Conversely, the entries with higher probability in N are the evaluative expressions that better represent the words.

The composition F then yields a similarity between words, not by their meaning (the words key and mouse are not closely related by meaning) but by their relation to the evaluative expressions: both key and mouse are small. Conversely, having the composition $G = NM$ yields an endomorphism of X en $\mathcal{P}_T$. Here the relation between two expressions is given by the closeness to their related words.

This can be pictured in the following diagrams:



In particular, this means that by looking at the entries of the compostion matrices we get a fuzzy Johari window decomposition of the concepts by looking at the probabñities of the matrices M, NF and G

4 Conclusions

To summarize, by constructing the semantic categories $\mathcal{L}_T$ and $\mathcal{P}_T$ we can infer the meaning of evaluative expressions from relative to subset of expressions of a text. This yields a probability distribution for each $g \in X$ of closely related elements of X (or of $\mathcal{L}_T^1$ in the case of M). The conclusions one can derive from these distributions is that the evaluative expression g is similar, i.e. applies to the same elements of $\mathcal{L}_T$, to $h \in X$ since their probability entry in G is close to 1.

To conclude this section, let us talk about possible computational implementations of these ideas. There are a few things to consider. The first and most basic is the text itself. It needs to contain as diverse as possible usages of the evaluative expressions we want to analyze. Once the text is selected, the next thing is to select a subset of elements A of $\mathcal{L}_T$ to compare the evaluative expressions with. This can be the whole $\mathcal{L}_T$ but it would make the algorithm quite slow.

The last thing to consider is how to compute the probabilities. Most of what is done in Natural Language Processing (NLP) tasks relies on these probabilities. The methods used to compute them often rely on counting co-occurrences of the expressions. Then, there are several normalization or smoothings that can be applied to the counts to avoid computing errors. All these choices will have an impact on the morphisms M and N and thus, the computed relations between the expressions are also depend on those choices.

References

1. Bradley, T.D., Terilla, J., Vlassopoulos, Y.: An enriched category theory of language: from syntax to semantics. La Matematica pp. 1–30 (2022)
2. Fritz, T., Gonda, T., Perrone, P., Rischel, E.F.: Representable markov categories and comparison of statistical experiments in categorical probability. Theoretical Computer Science **961**, 113896 (2023)
3. Perrone, P.: Markov categories and entropy. IEEE Transactions on Information Theory (2023)