INTERSPEECH 2023
20-24 August 2023, Dublin, Ireland

Cross-lingual/Cross-channel Intent Detection in Contact-Center Conversations

Suraj Agrawal[1]*, Aashraya Sachdeva*[1]*, Soumya Jain, Cijo George, Jithendra Vepa*

Observe.AI, India

{suraj.agrawal, aasharaya.sachdeva, soumya.jain, cijo.george, jithendra}@observe.ai

Abstract

Contact centers play a critical role in providing quality customer service for various businesses. Identification of business-related intents in these conversations is valuable for compliance, quality assurance, and analytics. This is often done using keyword spotting. But, identifying the right key phrases for accurate intent detection is challenging, especially for noisy call transcripts generated by ASR systems. Moreover, the process has to be repeated across different languages & to adapt existing intent definitions from other communication channels to call transcripts. We present a novel Semantic Phrase Search Engine that enables cross-lingual/cross-channel discovery of intent key phrases. The platform facilitates automated expansion of intent key phrases to improve coverage, and in-context translation of phrases across languages/channels, while maintaining semantic integrity. Thereby, our system helps in improving quality as well as reducing cost of contact center operations.

Index Terms: contact center, quality assurance, business intelligence, intent detection

1. Introduction

In contact centers, millions of customer calls are recorded, transcribed using ASR systems, and stored daily for use cases, including compliance, quality assurance, and analytics. These objectives often necessitate the identification of pertinent intents or "business-related events." To efficiently carry out this process, contact centers typically employ automated keyword spotting, where intents are characterized by a collection of key phrases and tracked throughout conversations.

However, the manual task of determining the most suitable key phrases for accurate intent identification in specific business contexts can be laborious and challenging due to the noisy nature of calls, and ASR errors. The process has to be repeated for different languages, requiring linguistic expertise since a direct translation of key phrases does not often yield optimal results. Moreover, the effectiveness of the same phrases may vary across different communication channels, such as calls, chat, or email, due to inherent discrepancies in phrase usage for expressing identical intents. For instance, "Thank you for calling" applies in call interactions, but the term "calling" does not appear in chat or email exchanges. This implies that intent definitions created for a different channel cannot be used for call transcripts. Consequently, contact centers must invest resources in developing intent representations tailored to each supported language and communication channel.

In this research, we introduce a system aimed at supporting human efforts in defining intent key phrases by suggesting phrases that are assured to be present in conversations in the target language/channel. Specifically, we present a Semantic Phrase Search Engine that supports cross-lingual/cross-channel phrase discovery. The system also supports the following capabilities:

- Intent Phrase Expansion - to improve coverage of intent detection by expanding the phrases list.
- Cross-lingual/cross-channel translation of intent definition - to enable the in-context translation of a set of intent key phrases to the target language and/or channel while maintaining semantic integrity.

[1]Equal contribution

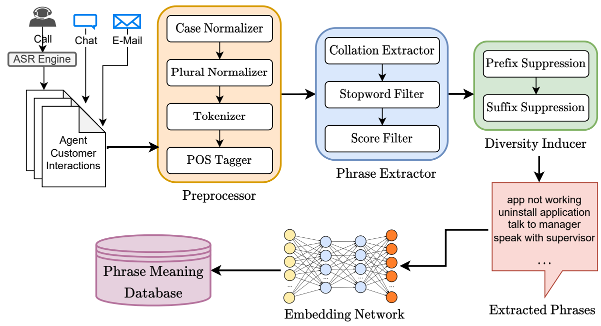

Figure 1: *Unsupervised Key-phrase Extraction Pipeline*

2. Key-phrase Mining Algorithm

The key-phrase mining algorithm is an unsupervised method used to extract the most relevant key-phrases from interactions between contact center agents and customers. It involves the following steps, as depicted in Figure 1:

1. Pre-processor: The pre-processor accepts input text (transcripts in the case of voice calls) and performs a series of language and channel-specific transformations.
   - Case Normalization: converts the text into a consistent lowercase format.
   - Plural Normalization: changes plural words into their singular form (e.g., "managers" to "manager").
   - Tokenizer: splits sentences into individual tokens.
   - POS tagging: assigns Part-Of-Speech (POS) tags to tokens using pre-built Spacy models.
2. Phrase Extractor: This module extracts relevant key-phrases from POS tagged tokens using following steps:
   - Collocation Extractor: Collocation is a series of words or terms that co-occur more often than would be expected by

chance. To find collocations, we first find all sequence of n-grams ($n \in 2, 3, 4$) that satisfy hand curated rules on POS tags.
- Stop-word Filter: The collocations are then filtered based on hand curated list of stop-words.
- Score Filter: Lastly, the collocations are filtered based on their statistical significance. We use Point-wise Mutual Information [1] to score each of collocation. Collections below a certain threshold are discarded.

3. Diversity Inducer: Many phrases extracted in step 2 have overlapping words amongst them. For example. “speak to manager”, “i speak to manager” have three words in common. Although, these are separate key-phrases, both of them convey the similarly meaning. The prevent duplicate prediction, we iteratively merge smaller key-phrase into a bigger key-phrase if they have a common prefix (or suffix) and they both have almost same number of occurrences.
4. Embedding network: Each extracted key phrase is then encoded into a 512-d vector. We use Google’s Multilingual Universal Sentence Encoder [2] as our embedding network, primarily due to its ability to support out-of-vocab words and multilingual support.
5. Phrase-meaning Database: Finally, pairs of key-phrases and their vectors are stored in a database.

3. Intent Phrase Expansion

Intent phrase expansion is a nearest-vector search algorithm that, given an example query phrase (which defines an intent), suggests alternate phrases which mean the same thing, but also present in interactions. The overall workflow is in Figure 2.

1. User enters a phrase (query) defining an intent.
2. The query goes through a basic pre-processing module viz-viz case and plural normalization.
3. The pre-processed query is then converted to a 512-d vector using the embedding network as defined in Section 2.
4. A set of phrases having similar meaning are extracted from the database using a cosine distance metric. We call these phrases, semantic predictions.

Semantic predictions corresponding to a query can then be used to define or extend an intent. Since a personalized database is curated for every contact center, the predictions are assured to increase coverage.

4. Cross-lingual/Cross-channel Translation

Given a list of phrases defining an intent in the source language/channel, this module provides a list of semantically translated phrases in the target language/channel. As before, semantically translated phrases are assured to be present in the interactions. The algorithm is described in the following steps:

1. First, we run the key phrase mining algorithm to extract the relevant key phrases from the interactions present in the target language.
2. Then, we derive the top-k semantically similar phrases, for each query phrase by running the intent phrase expansion algorithm.
3. Finally, we merge the list of candidate phrases obtained for every query phrase and rank them based on weighted scores. Phrases are then filtered based on a weighted score threshold. The weighted score for each candidate phrase p is calculated

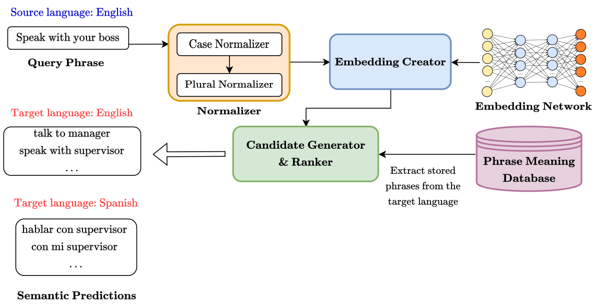

Figure 2: *Phrase expansion pipeline*

as follows:

$$score_p = score_{1p} + \frac{score_{2p}}{2} + \frac{score_{3p}}{3} + \ldots \quad (1)$$

where, $score_{ip}$, is the similarity score between query phrase and candidate phrase p.

5. Data Liveliness

In the context of a contact center, where interactions occur continuously, one may question how the Phrase-Meaning Database remains up-to-date with new interactions. The solution is straightforward: the key-phrase mining algorithm is executed at regular time intervals. As speaking trends change, modifications in key-phrases are updated within the database. Consequently, the phrase suggestion pipeline predicts key-phrases that mirror current trends, ensuring the intent definition aligns with these trends. A prime example of this dynamic nature can be observed in the intent ”Did the agent provide a meaningful closing to the consumer when ending the call.” Traditional key-phrases for this intent include ”have a great day,” ”goodbye,” and ”have a good week ahead.” Nevertheless, during the COVID-19 pandemic, a shift in trend emerged. Thanks to the data liveliness, the algorithm predicted new key-phrases for this intent, such as ”take care” and ”stay safe.”

6. Conclusion

This work addresses the challenge of identifying appropriate key phrases for intent identification in contact centers with multiple languages and communication channels. The Semantic Phrase Search Engine proposed in this study offers a practical solution by suggesting phrases that are certain to appear in the target language/channel. The system can improve the precision and efficiency of intent identification in contact centers while reducing manual efforts significantly.

7. References

[1] G. J. Bouma, “Normalized (pointwise) mutual information in collocation extraction,” 2009.

[2] Y. Yang, D. Cer, A. Ahmad, M. Guo, J. Law, N. Constant, G. Hernandez Abrego, S. Yuan, C. Tar, Y.-h. Sung, B. Strope, and R. Kurzweil, “Multilingual universal sentence encoder for semantic retrieval,” in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*. Online: Association for Computational Linguistics, Jul. 2020, pp. 87–94. [Online]. Available: https://aclanthology.org/2020.acl-demos.12