📚➡️🙂

An Approach for Text-to-Emoji Translation

Philipp Wicke
Creative Language Systems
University College Dublin
Belfield 4, Dublin - Ireland
philipp.wicke@ucdconnect.ie

João M. Cunha
CISUC, Dept. of Informatics Engineering
University of Coimbra
3030-790 Coimbra - Portugal
jmacunha@dei.uc.pt

Abstract

The task of translating text to images holds some valid creative potential and has been the subject of study in Computational Creativity. In this paper, we present preliminary work focused on emoji translation. The work-in-progress system is based on techniques of information retrieval. We compare the performance of our system with three deep learning approaches using a text-to-emoji task. The preliminary results suggest some advantages of using a knowledge-base as opposed to a purely data-driven approach. This paper aims to situate the research, underline its relevance and attract valuable feedback for future developments.

Introduction

Computational approaches to automatic illustration of text have long been a subject of study. Several methods have been explored to produce visual representations in the field of Computational Creativity. Some consist of systems for collage generation, e.g. Cook and Colton (2011), others employ visual blending techniques, e.g. Xiao and Linkola (2015), and generative adversarial networks (GANs) have also been used to synthesise photorealistic images that represent a given theme, e.g. Ni et al. (2020).

When it comes to using images to convey concepts or ideas, it is impossible not to mention emoji. According to a report[1] released by Adobe in 2019, emoji continue to thrive – users surveyed admit to include emoji in text messages 49% of the time. Despite being mostly used as a way to make conversations more fun or lighten the mood, 94% of the surveyed users identify the ability to communicate across language barriers as one of the greatest benefits of emoji. Two different functions of emoji can be identified (Dürscheid and Siever, 2017): complementary (to accompany text) and replacement (to replace words). While the former has been often computationally explored for the development of emoji prediction methods, the latter has not been given the same attention. An exception is the Emoji Replacement function introduced in iOS 10[2]. Nonetheless, this exception mostly addresses the replacement of single words.

[1] www.slideshare.net/adobe/adobe-emoji-trend-report-2019/1

[2] https://macrumors.com/how-to/ios-10-messages-emoji/

In this paper, we present an approach for text to emoji translation. In order to build such a system there are many different avenues to pursue. In the day and age of *Deep Learning*, the most obvious technique would be a data-driven approach using Machine Learning (ML). An issue with this approach is that no large corpus of text translated to emoji exists – even though emoji are highly used, translations of entire texts to emoji are rare. An exception is *Emoji Dick* by Benenson (2010), which has been crafted using human crowd-workers – an exploratory analysis by Radford et al. (2016) indicated that semantics were preserved in the translation. As a consequence, ML systems that have been trained entirely with correlations of co-occurrences of words and emoji often fail to produce viable text-to-emoji translations. On the one hand, this is obvious from a linguistic perspective as emoji do not provide a syntax and are not designed to substitute a language (Cohn, Engelen, and Schilperoord, 2019). On the other hand, the need to incorporate emoji into NLP applications is vast (Xu et al., 2018; Eisner et al., 2016).

The system that we proposed is not based on ML techniques, but on information retrieval techniques and knowledge bases in order to overcome the issue that data-driven approaches are facing. Our system is inspired by insights in emoji and symbol research from various other research studies Wicke (2017); Wicke and Bolognesi (2020); Cunha, Martins, and Machado (2018). In this paper, we test our system and three ML-driven systems in a text to emoji translation task. In the methods section, we briefly present the key mechanisms of all four systems before we present the preliminary study on emoji translation. As this paper presents work in progress, the results of the systems are at this point evaluated and discussed qualitatively by the authors. In future work, we will describe how the results can inform an empirical evaluation that will eventually improve not only our own translation system but also the understanding of emoji in language.

Related Work

Several related works have inspired the methods applied in the proposed translation system. Closely related to our text-to-emoji system is the one presented by Wicke (2017). The author creates and evaluates a system that can translate action words into sequences of emoji through the use of vari-

Proceedings of the 11th International Conference on Computational Creativity (ICCC'20)
ISBN: 978-989-54160-2-8

ous linguistic strategies (metaphor, idioms, rebus etc). In the empirical evaluation the author concludes that action word translations using the rebus principle (e.g. 🐝 🍁 for "believe"), metaphors (e.g. ☘ for "luck") or literal translations (e.g. 💥 for the action "to explode") are being best understood and appreciated by human readers.

If our system is tasked to translate text to emoji, we also consider how humans are performing in such a task. In a study by Wicke and Bolognesi (2020), the authors had 300 concept words (e.g. dog, democracy, luck, family) translated by crowd-workers using *Amazon Mechanical Turk*. With 11 translations per concept from individual workers, the authors analysed the coherence of translations correlated to the concreteness and abstractness of each concept. The results of their study indicate the more concrete the concept, the greater the coherence and the more abstract the concept, the more emoji are being used and the more face emoji occur.

Another approach to concept representation using emoji is *Emojinating* (Cunha et al., 2019) – a system that uses visual blending of emoji to represent user introduced concepts. However, this system is of little use for our study as it considers one-word and two-word concepts.

Regarding text to emoji translation, we describe three ML-driven systems that we will compare our system against: SemEval System, DeepMoji and DangoApp.

SemEval System is described by Çöltekin and Rama (2018) and was a contribution at 2018's International Workshop on Semantic Evaluation Task 2: Multilingual Emoji Prediction. This model is based on a One-Vs-Rest Support Vector Machine architecture and has been trained on 500,000 Twitter messages (tweets). The training input was one tweet to predict one emoji on the basis of 20 classes of emoji (the most common emoji on Twitter at that time). The requirement for one tweet one emoji prediction was given by the SemEval task and is an obvious restriction in a text-to-emoji translation. Yet, this system has been included in the study as the code is freely available[3], easy to implement and allows us to reproduce the model.

DeepMoji is a system for the detection of sentiment, emotion and sarcasm in text using emoji, implemented by Felbo et al. (2017). *DeepMoji* is also freely available[4]. This system is trained on 1.2 billion tweets containing one of the 64 most common emoji on Twitter. The neural network architecture is comprised of: Embedding Layer → BiLSTM → BiLSTM → Attention Layer → Softmax. Again, this project does not aim to translate text-to-emoji directly but to reflect the emotional content of a tweet through emoji.

DangoApp The fourth system is Software Product by *WHIRLSCAPE* released for Android/iPhone as an App. The self-titled "*Emoji Assistant*" is a real-time emoji prediction app. It claims to also capture slang expressions and memes as a result of its *Deep Learning* architecture with RNNs providing a semantic space for cosine similarity measures. All information is provided on their website: *getdango.com*. As the code and the model are not available, we use the app to test its text-to-emoji translation capabilities.

[3]https://github.com/coltekin/emoji2018

[4]https://github.com/bfelbo/DeepMoji

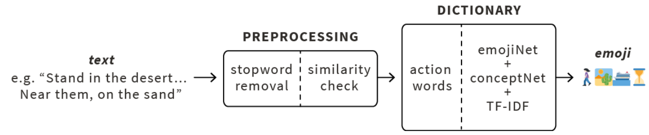

Figure 1: Overview *InfoRet* System

Our Approach: InfoRet

The goal of our system is the translation of a sequence of words (text) to a sequence of emoji that can convey the same or similar meaning. Contrary to all the presented ML approaches use data that has text-with-emojis instead of text-to-emoji, we decided to distance our approach from those. Instead, we adopted the insights from Wicke (2017) and used the database of action-to-emoji mappings in combination with ConceptNet (Liu and Singh, 2004) and EmojiNet (Wijeratne et al., 2017). Our method is explained in the next subsection and as those techniques are mostly information retrieval techniques, we call the system hereinafter *InfoRet* (Fig. 1).

Dictionary Creation In its essence, the proposed system is built around a dictionary of translations from words to emoji. The core part of the research is the constitution of this dictionary. It includes the following entries:

1. The 400+ action-to-emoji translations by Wicke (2017).
2. All entries from *EmojiNet* Wijeratne et al. (2017) – a machine-readable emoji sense inventory that maps Unicode emoji representations to their English meanings extracted. It consists of 12,904 sense labels for over 2,389 emoji. Within the labels, we perform a term-frequency inverse document frequency (tf-idf) analysis to weight the most important emoji for each label. We access ConceptNet in order to extend the labels provided. For example, the label "dog" can be extended with "canine" or "puppy" using ConceptNet.

The idea behind the first addition of action-to-emoji translations has been explained previously, the second addition of *EmojiNet* needs a brief explanation. With 12,904 sense labels, we are provided with a great addition to the dictionary. Yet, the emoji that are linked to a label by the Unicode are often very similar, e.g. "animal" is a label attached to every emoji that depicts an animal. If we would just use the first label, we would lose important information such as "animal, dog". Performing a tf-idf on those labels allows us to extract the label most relevant to its overall frequency. Now, we can greatly expand the labels if the label "dog" can also refer to "puppy" or "canine" using ConceptNet.

Translation The system takes a sentence as an input. The sentence will be filtered for common stopwords (I, me, am, him, his etc) (Stone, Dennis, and Kwantes, 2011). For each word in the sentence, it is checked whether the word is similar (similarity checked here using Python's difflib SequenceMatcher) to an entry in the dictionary. If there is a match the corresponding emoji will be stored.

Figure 2: Results of translations by the four systems. A: SemEval, B: InfoRet, C: DeepMoji, D: DangoApp

.

Study on Emoji Translation

In this section, we describe a preliminary study that tests our approach and three ML-driven systems in a text to emoji translation task.

Experimental Setup

We compare the four approaches in a text-to-emoji task. Even though our ultimate goal is to translate text to emoji in short stories, we decided to test the systems with other types of text as well. Therefore, we picked three different types of texts: a tweet (as this is what most of the ML approaches have been trained on), a short story (as this would be the ideal text length to be translated) and a poem (to compare creative aspects in a different domain). The tweet is one by Donald Trump and one by Barrack Obama, the short story is "Appointment in Samarra" by W. Somerset Maugham and the poem is "Ozymandias" by Percy Bysshe Shelley. We ran each of the four systems on each of the four texts (sentence by sentence).

Results and Discussion

In Fig. 2 we can see the results of each system for the respective text. We will interpret the results for each system separately before we conclude with a comparison.

SemEval: This method is primarily focused on the labelling of one tweet with one emoji. We can see the limitations for translating full text in the results i.e. there are only three individual emoji used for the translation: 26x😂, 3x💯, 2x😊. Even for its target domain (tweets), the system annotates both tweets with the same "tears of joy"-emoji, despite the fact that both tweets carry a different sentiment. As this approach seems to fail in its own domain, there is hardly any use in other textual domains such as short story or poems. As to be expected, the results indicate that a sentiment classification with only one label per sentence is too far from the system we need for text-to-emoji translation.

InfoRet: The results of the InfoRet system are the most diverse and will need the most interpretation. As we can see in Fig. 2, the system generates sequences between zero (Short story line 10, poem lines 9 and 13) and six (Trump tweet line 2, poem line 6) emoji long. This is a variability we do not observe for the other systems. The cases of zero emoji can be considered a failure of the system, which occurred three times. This also means that the system "keeps quiet" when there is no definite solution, something the other systems do not account for. We can observe some instances in which the InfoRet system suggests a much better translation than the other systems. For example in line 10 of the poem, the InfoRet suggests the "name badge" and "crown", whereas only the DangoApp captures the word "King", but not the naming aspect of the text. We can also see reoccurring tweets over multiple lines, e.g. the button as defined in Trump's tweet's translation as 🔽. This button can be observed over multiple tweets, which is a feature that the other algorithms do not provide. This might suggest, that InfoRet is more useful for long text translations.

DeepMoji: This method also does not aim to be a text-to-emoji translation but rather to express the sentiment of a sentence. Therefore, it is useful for sentiment analysis of a sentence, but not necessarily for a direct translation. We can observe a strong overrepresentation of the music, keys and notes emoji in the poem, even though none of this is related to the text. Investigating the Trump tweet we can see that the first line is comprised of tweets signalling the same ridicule sentiment as the tweet, whereas the second line presents the negative sentiment of the text and the last sentence reflects self-praise with emoji. As good as this system captures the sentiment as bad can you infer any meaning of the underlying text.

DangoApp: The results of the DangoApp can be seen as somewhat in between the InfoRet and the DeepMoji. It captures sentiment, yet it also allows including more concrete emoji. We can compare some examples: The Donald Trump tweet includes (line 1) the ridicule with the laughing emoji, but it also includes a flag (not the North Korean one though). Line 3 shows the positive face emojis and a flash emoji to relate to the "powerful" in the text. Yet, we can also see how this system fails in the poem line 9, where the musical emoji do not relate to the text.

Overall, our symbolic/information retrieval approach seems to show advantages over the machine learning approaches. The three deep learning approaches show strong deficits as soon as we leave the Twitter domain, due to the fact that they were trained on Twitter data. In fact, comparing these approaches in an emoji translation task of longer texts can be considered unfair as they have not been built for this purpose. Nonetheless, as there is no dataset of stories, poems or longer texts that are translated to emojis – and it is unlikely such a database will soon be created – the InfoRet is most likely the best approach to serve as a domain-unspecific, general model for translating text into emoji. Our system does not seem to have much use for communicative purposes, as using emoji for word-replacement is not very appropriate for written communication and has even been shown to increase reading time (Gustafsson, 2017). For artistic purposes, we consider that InfoRet has potential, especially when it comes to automatic text illustration.

Conclusion and Future Work

In this paper, we presented an approach for automatic text to emoji translation, based on information retrieval techniques. We tested the system with three different types of texts: tweet, short story and poem. We compare the results with the ones from three machine learning-based systems. Notably, our evaluation is of preliminary, subjective nature. This paper describes work in progress and, as such, a more thorough validation needs to be conducted. Notably, a similar approach was made public after we conducted our evaluation in Day et al. (2020 forthcoming). For future work, it will be highly valuable to compare our approach with this system once it has been published.

Acknowledgements This work is partially funded by the FCT - Foundation for Science and Technology, I.P., within the scope of the project CISUC - UID/CEC/00326/2020 and by European Social Fund, through the Regional Operational Program Centro 2020, and under the grant SFRH/BD/120905/2016.

References

Benenson, F. 2010. Emoji dick. *Creative Commons Attribution*.

Cohn, N.; Engelen, J.; and Schilperoord, J. 2019. The grammar of emoji? constraints on communicative pictorial sequencing. *Cognitive research: principles and implications* 4(1):33.

Çöltekin, Ç., and Rama, T. 2018. Tübingen-oslo at semeval-2018 task 2: Svms perform better than rnns in emoji prediction. In *Proc. of SemEval-18*, 34–38.

Cook, M., and Colton, S. 2011. Automated collage generation-with more intent. In *ICCC*, 1–3. Citeseer.

Cunha, J. M.; Lourenço, N.; Correia, J.; Martins, P.; and Machado, P. 2019. Emojinating: Evolving emoji blends. In *Proc. of EvoMusArt 2019*. Springer.

Cunha, J. M.; Martins, P.; and Machado, P. 2018. Emojinating: Representing concepts using emoji. In *Workshop Proceedings from ICCBR*.

Day, A.; Mankos, C.; Kim, S.; and Strausser, J. "2020, forthcoming". Confet: An english sentence to emojis translation algorithm. In *Proceedings of the 35th Annual Spring Conference of the Pennsylvania Computer and Information Science Educators*.

Dürscheid, C., and Siever, C. M. 2017. Beyond the alphabet–communication of emojis. *Kurzfassung eines (auf Deutsch) zur Publikation eingereichten Manuskripts*.

Eisner, B.; Rocktäschel, T.; Augenstein, I.; Bošnjak, M.; and Riedel, S. 2016. emoji2vec: Learning emoji representations from their description. In *Proc. of SocialNLP 2016*.

Felbo, B.; Mislove, A.; Søgaard, A.; Rahwan, I.; and Lehmann, S. 2017. Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion and sarcasm. In *Proc. of EMNLP*.

Gustafsson, V. 2017. Replacing words with emojis and its effect on reading time. *USCCS 2017* 73.

Liu, H., and Singh, P. 2004. Conceptnet-a practical commonsense reasoning tool-kit. *BT technology journal*.

Ni, J.; Zhang, S.; Zhou, Z.; Hou, J.; and Gao, F. 2020. Instance mask embedding and attribute-adaptive generative adversarial network for text-to-image synthesis. *IEEE Access* 8:37697–37711.

Radford, W.; Chisholm, A.; Hachey, B.; and Han, B. 2016. : telephone:: person:: sailboat:: whale:: okhand:; or" call me ishmael"-how do you translate emoji? *arXiv:1611.02027*.

Stone, B.; Dennis, S.; and Kwantes, P. J. 2011. Comparing methods for single paragraph similarity analysis. *Topics in Cognitive Science* 3(1):92–122.

Wicke, P., and Bolognesi, M. 2020. Emoji-based semantic representations for abstract and concrete concepts. *Cognitive processing*.

Wicke, P. 2017. Ideograms as semantic primes: Emoji in computational linguistic creativity.

Wijeratne, S.; Balasuriya, L.; Sheth, A.; and Doran, D. 2017. Emojinet: An open service and api for emoji sense discovery. In *Proc. of ICWSM-17*.

Xiao, P., and Linkola, S. 2015. Vismantic: Meaning-making with images. In *Proc. of ICCC-15*.

Xu, P.; Madotto, A.; Wu, C.-S.; Park, J. H.; and Fung, P. 2018. Emo2vec: Learning generalized emotion representation by multi-task training. *arXiv:1809.04505*.