Emolift: Elevator Conversations Based on Emotions

Alejandro Oñate[1] and **Gonzalo Méndez**[1,2] and **Pablo Gervás**[1,2]
[1]Facultad de Informática
[2]Instituto de Tecnología del Conocimiento
Universidad Complutense de Madrid
{aleonate,gmendez,pgervas}@ucm.es

Abstract

Most of the currently existing narrative generation systems do not consider the task of adding dialogues to enrich the interactions between characters. An important factor to take into account when two characters speak are the emotions they feel and the ones they perceive from the character they are talking to. In this work we present Emolift, an automatic dialogue generation system of short elevator dialogues where emotions are the driving force to build dialogues. We have implemented a system capable of creating dialogues based on the emotions of the characters and the perception of other character's emotions in order to nourish and enrich storytelling systems.

Introduction

Finding something interesting to say to the person standing next to you in the lift is a challenge of everyday life that we all face frequently. Conventional openings like the weather can fill in the awkward silence, but more creative solutions are to be preferred. The present paper explores the assumption that the individual contributions of an agent initiating or participating in such conversations may be generated by assuming a simple underlying purpose generally not related to the actual content under discussion of equalizing the emotional states of the participants in the conversation.

The purpose of this work is to develop an automated dialogue generation system in Spanish where the focus is set on the emotions involved in the conversation. With this objective in mind, we have chosen elevator dialogues, where we usually have conversations without relevant content and with people with little or no affinity. In this way, the variable that most affects the dialogue is the emotional state of the characters. The final goal of this task is to create a component that can be used in storytelling systems to enrich the results produced by traditional generators, which in general tend to neglect this part of the stories. Therefore, for any situation reported in indirect speech, a direct speech equivalent can be generated to provide a better picture about how the characters of the story acted in that situation.

In (Minsky 2010) the author stated that emotions serve as regulators of behavior, since intelligence is characterized, among other aspects, by being able to adjust to a situation. A system should have different outputs for the same entries depending on the involved emotions. For example, we tend to be more tolerant of the people to whom we have more affection (or affinity), just as our answers generally depend on the way we have received a question.

Therefore, for a conversation to be as realistic as possible, the machine must take the emotional factors into account. Otherwise, if we generate dialogues without considering these parameters, the conversation may be too linear or, on the contrary, give emotional jumps creating totally unusual responses in a real human conversation.

In many storytelling systems, the ability to generate dialogues between characters is still not present (Cavazza and Charles 2005) and practically none of them incorporate emotional factors or characters' personality. Therefore, our most ambitious goal is to create a dialogue generation component for storytelling systems where we can solve this issue: to use the affinity, personality and emotions of the characters with the aim of improving the outputs created by the story generators by adding emotional dialogues between their characters.

In our previous work related to automatic dialogue generation, we approached this problem by using two of the dialogue's modifying parameters: affinity and mood (Oñate, Méndez, and Gervás 2019). Our goal was to be able to generate brief dialogues with a question-answer style to decorate the plan proposals generated by the Charade storytelling system (Méndez, Gervás, and León 2016).

To carry out this task, we take into account the current emotions of the character (base emotions) and the emotions that the character perceives about the speaker (external emotions) with the aim of developing a system capable of creating short conversations (machine-to-machine) like the one that two humans could have in an elevator, considering the emotional values of each character and the impact that the emotions of each character have on the emotional state of the other. This system is called **Emolift**.

In the next sections we present some related work on emotional dialogue generation, followed by a detailed description of the Emolift system accompanied by some examples of its behaviour. We finally present the conclusions of our work, together with the future lines of work to develop a more complete automatic dialogue generator for storytelling systems.

Related Work

Since our goal is to implement a system capable of generating dialogues adapted to the emotions that people feel and the emotions they perceive from their interlocutor, it is necessary to make a brief review of the three phases that our system will face. First, the generation of dialogues, then the analysis of emotions from a text and finally the generation based on emotions.

Dialogue generation

Traum (2017) states that dialogue management systems do not offer a novel interaction. These systems generate dialogues obtained from the analysis of human data, instead of allowing the systems to develop new comprehension and capabilities through the use of dialogues.

In (Oh and Rudnicky 2000) a new approach to generate language for systems of spoken dialogue is presented. For the surface realization task, the authors use an n-gram language model to stochastically generate each expression and observe that this stochastic system works at least as well as the template-based systems. This stochastic generation mechanism has several advantages over template-based ones, such as the response time. Another advantage they point out is that, by using a corpus-based approach, they are directly imitating the language of an expert in real domains, instead of trying to model the sentence generation as a rule, which the authors highlight as a positive factor but that contradicts the conclusions presented in (Traum 2017).

Apparently, this is a very recurring problem, as the Loebner Prize Competition has been used to evaluate the ability of chatbots to fool people making them believe that they are speaking to humans. The chatbots learn to imitate a human using expressions and terminologies that a human would use, while the real objective should be to perform the dialogue generation task for which it has been programmed in the best possible way (Shawar and Atwell 2007).

Emotional analysis

In order to be able to imitate the emotions of humans, it is necessary to analyze how they express themselves in different situations. Therefore, we can see how in the last decade there has been an exponential growth in the field of emotional analysis, and it can be reflected in the number of academic articles published (Lerner et al. 2015).

Thanks to the volumes of information that we expose daily in a public way in social networks, using all kinds of languages (formal, technical, colloquial) and on various subjects, it has been possible to create a wide corpus of emotional information. A clear example is (Pak and Paroubek 2010), a reference when training Neural Networks for the classification of texts according to three types of polarities: positive, neutral or negative.

This type of analysis has quickly spread and led to more specialized studies, since it is no longer enough to know whether a text is positive or negative, but we also need to interpret the emotions that a writer reflects in his/her texts. In this way, researchers began to model mood, subjectivity and emotion detection systems in texts published on social networks such as (Wang et al. 2012). Even more, some systems use these techniques to combine aspects like the force indicators of opinion, emotion and polarity in order to obtain analysis significantly more accurate than the previous ones (Bollen, Mao, and Pepe 2011).

According to the cues for emotion expression, there are two main methods for sentence emotion recognition: *emotion provoking event* based method and *emotion words* based method. Quan and Ren (2010) use the emotion words based method, which is seen as the most naive approach but also the most popular one. This method consists in extracting the emotional value of a sentence by analyzing each of the words that comprise it. This is the technique we have used throughout this work, which has already been tested for analyzing text in Spanish (Miranda, Luna, and Morillo 2016) and in automatic narrative generation (Delatorre et al. 2017).

Dialogue generation based on emotions

Having analyzed the generation of dialogues and the analysis of emotions in texts, it is time to see where the use of emotions can enrich the generation of texts in general, and the generation of dialogues in particular.

An example where the generation of text based on emotions helped empathize with people is (Mahamood and Reiter 2011). This work uses several Natural Language Generation (NLG) strategies to produce affective medical reports for parents of neonatal babies undergoing medical care. The authors of this work report that all recipients preferred texts generated with affective strategies, regardless of the expected level of stress.

In (Ghosh et al. 2017) a new Affect-LM language model is presented to generate affective conversational texts, which are conditioned by context words, an affective category and a parameter of affective force. This study shows that the model can generate expressive text with varying emotional strength without affecting the grammatical correctness.

With the arrival of personal assistants, research on chatbots and Embodied Conversational Agents (ECA) has intensified in recent years. The most common application areas of the ECAs are education, entertainment, search engines, customer service and support. Polzin and Waibel (2000) present how dialogue behaviors can adjust to the emotional states of users, so it is becoming essential to add interpretation and generation based on emotions to this type of systems.

Finally, and as the greatest success in the generation of emotional dialogues, Zhou et al. (2018) have managed to create a question-answer style system taking into account the five basic emotions that they have considered: Like, Happy, Sad, Disgust and Angry. To carry out this Emotional Chatting Machine, the authors use Fuzzy Logic to discretize the emotional values of the sentences, and offer a response of the emotion obtained. However, this work is only a question-answering system and it is not capable of generating an extensive dialogue.

Emolift System

As we have previously explained, our goal is to elaborate a system capable of generating a dialogue between two characters that meet in an elevator, as a simplified situation that will later be extrapolated to contextualized dialogues in a story. In this dialogue, we will take into consideration the current state of the emotions felt by each character when they are talking to each other. These emotions are affected by the way our companion addresses us (i.e. the emotions we perceive from their words) (Polzin and Waibel 2000).

It is important to remember that, by taking place in an elevator, these conversations tend to have a peculiar context, and many times they will turn out to be empty conversations, led just by compromise, apart from being really short conversations of approximately 20 seconds long. Because of that, it is not necessary to keep track of emotions, since the conversation takes place in a very short period of time.

This system consists of three fundamental stages. First, the analysis of the emotions transmitted by the sentences, through which we can classify and analyze the sentences of the dialogue. Second, the emotions updater, which will help us modify the base emotions of the characters, combining them with the emotions perceived from every sentence we hear. And, finally, the generation of text based on our updated base emotions, to provide an answer according to our current emotional state. We have addressed these three issues, to have more control and information about what is happening within the conversation. Although the focus of this work is to be found on the last stage of the system, the generation of sentences with emotional value, the three stages are fundamental to be able to generate dialogues based on emotions.

Emotional analysis of a sentence

Since our main goal is the generation of dialogues, and not the analysis of texts itself, and given that high quality data tagged with emotions is hard to obtain as a large-scale corpus, and the annotation of emotions is a very subjective task (Ghosh et al. 2017), we have decided to create a simple system that acts as an analyzer, with the goal of being able to classify the sentences of our dialogue generator.

At the beginning of this work, we made use of two affective dictionaries collecting a total of 3,141 words in Spanish. The first of them (Ferré et al. 2017) consists of 2,266 words, all of them tagged with the same five emotional categories we need. The second dictionary (Hinojosa et al. 2016) includes affective norms for 875 words included in MADS (Madrid Active Database for Spanish) and we chose it because the purpose of this dictionary is to complement the corpus of (Ferré et al. 2017), our main dictionary. As this work was being carried out, a new research was published by the same authors (Stadthagen-González et al. 2018) providing a new version of the affective dictionary that contains three times more words than the other two combined (10,491). By just loading these new words and their values in our system, it was enough to improve its analysis and emotional generation.

The words included in this last dictionary have a value from 1 to 5 for each of the considered emotions according to the intensity with which it transmits that emotion. These emotions are joy, anger, sadness, fear and disgust.

With our dictionary already indexed, the task of analyzing existing emotions in sentences begins. For this, we have started using the mechanism described in (Eugercios, Gutierrez, and Kaloyanova 2018) for the development of EmoTraductor. We start from the basis that any sentence is devoid of emotion so, if it does not contain any emotional word, the result will be that it is a 100% neutral sentence. Therefore, the emotion of a sentence is based on the combination of the emotions of the words that make up that sentence. So, to analyze the sentence, we must previously analyze each of its words. To carry out this analysis, we will first identify the type of word for each of them.

The relevant words are the names, verbs, adjectives and adverbs. We search these words in our emotional dictionary with the objective of obtaining the emotions that they transmit. Since words have derivations, if the exact word is not found in the dictionary, we compare its lemmatization with the lemmatization of the dictionary words. In the latter case the same search is performed using the stemmer of the word.

Example in our dictionary:
Word: aceptamos (we accept)
Lemmatization: aceptar (infinitive)
Stemmer: acept (remove sufix)

Once we have the emotions for each of the words, we proceed to calculate the emotion of the sentence. Eugercios, Gutierrez, and Kaloyanova propose a calculation based on the arithmetic average of the emotions per word, in such a way that the emotion of the sentence *"Me alegro mucho"* (I am very happy) is calculated by adding the values of the words *alegro* and *mucho* and dividing by three.

This way of calculating emotions is functionally valid, but, from our point of view, it has a drawback. If we perform an arithmetic average, adding extra words with joyful emotion to our sentence will diminish the intensity of that emotion in the sentence. For a human being, naturally, the more you decorate a sentence with happy words or joyful adjectives, the happier it is. For example, if instead of *"Me alegro mucho"* (I'm very happy) we say *"Genial, me alegro mucho"* (Great, I'm very happy) or even *"Genial, me alegro mucho, me hace feliz saberlo"* (Great, I'm very happy, it makes me happy to know it), it is obvious that each of the examples conveys more joy than the previous one. Following the norm of the arithmetic average, less happy words would reduce the final value of the sentence.

Das and Bandyopadhyay (2009) propose a more complex formula, taking into account the value of all the emotions for the calculation of each of them. The value of each emotion is the total value of that emotion (STW_i) multiplied by the words that produce that emotion (N_i) divided by the sum of the values of each emotion (STW_j) and multiplied by their appearances (N_j):

$$SWS_i = \frac{STW_i * N_i}{\sum_{j=1}^{5} STW_j * N_j}$$

Evaluating this formula, we can see that the decorating words do not increase the emotion. We propose to find the

Spanish Sentence *(English)*	Ours	Avg.
Me alegro por ti *(I'm happy for you)*	3.69	2.12
Me alegro **mucho** por ti *(I'm **so** happy for you)*	3.74	2.01
Me alegro **mucho mucho** por ti *(I'm **so very** happy for you)*	3.78	1.92

Table 1: Comparison of our formula with (Eugercios, Gutierrez, and Kaloyanova 2018) to calculate *joy*

word with the highest value for each of the emotions, and use it as the base emotion for the whole sentence. For each of the remaining words, we take their dominating emotion and use its value to increment the base value of that emotion. We have determined empirically to apply an increase of 10% of its value, always respecting the top limit of 5 points. This way, we get the previous example as follows:

$$S_i = West_i + \sum_{k=1}^{n} Wres_{ki} * 0.1$$

where S_i is the "Current emotional value of the sentence", $West_i$ is the "Value of the word with more emotion" and $Wres_{ki}$ is the "Emotional values of the Residuary word in position *k*".

In **Table 1** we can see how our formula gives greater value to decorated sentences, where a human expresses greater joy.

Emotional dialogue generation

Once we have our emotional dictionary, and having defined the procedure of emotional analysis for sentences, we can start generating dialogues based on the emotions of the characters. To carry out this task, we have opted for a system based on templates, which is more controlled and allows us to evaluate and put the emotional analysis into practice.

Our goal is to create a dialogue generation system, not an emotional chatbot, so there is no intervention of a user as an external agent. This point is crucial, since our knowledge base must be more concrete and concise. This allows us to focus on clear objectives and therefore it is not necessary to prevent atypical inputs in the system.

For this reason, we have ruled out conventional template systems such as AIML[1] or RiveScript[2], and we have developed our own template system, where the input recognition does not prevail but the emotional variability of the output.

Each sentence of our templates makes use of a nomenclature that allows us to add alternatives when using words or expressions that generate emotions. The system will choose the most appropriate alternative in order to get the sentence that best suits the current emotion of the character.

Following the example of the previous section we could create a template like this:

Original Template
<|Genial,|Fantástico,|Estupendo>me alegro <por ti |mucho ><|, me hace feliz saberlo>

Translated Template
<|Cool,|Fantastic,|Great>I'm glad <for you |really ><|, it makes me happy to know it>

To make the selection of the sentence that has the most emotional similarity with the current emotional state of the character, the system analyzes each section (included in [ ]) and obtains the alternatives delimited by the separator (the character —). Note that the first and last sections begin with this separator, which allows us to indicate that among the available alternatives there is the option of not selecting any expression (select empty string).

Once we have identified each of the sections and their options, we generate all the possible combinations of sentences using the Cartesian product. This allows us to obtain a great variety of responses with different emotions for a simple sentence like this one. The options of this example and the emotional value of each of them would be:

Me alegro por ti
joy: 3.69, anger: 1, sadness: 1, fear: 1, disgust: 1

Me alegro por ti, me hace feliz saberlo
joy: 4.83, anger: 1, sadness: 1, fear: 1, disgust: 1

Genial, me alegro mucho, me hace feliz saberlo
joy: 4.84, anger: 1, sadness: 1, fear: 1, disgust: 1

Fantástico, me alegro por ti, me hace feliz saberlo
joy: 4.87, anger: 1, sadness: 1, fear: 1, disgust: 1

Fantástico, me alegro mucho
joy: 4.51, anger: 1, sadness: 1, fear: 1, disgust: 1

Estupendo me alegro mucho
joy: 4.32, anger: 1, sadness: 1, fear: 1, disgust: 1

Following this process, we have developed a corpus of brief dialogues typical of an elevator conversation. For each of them, we have added variations and alternatives like those shown in the previous example.

From now on, we will use an example of a real dialogue, in which two neighbors are in the elevator and one asks the other how he/she is. Among the possible answers, we have the following template, with which, after applying the Cartesian product, we obtain 1,008 possible combinations, with 878 different emotional combinations. This is because some combinations produce exactly the same emotion, since there are words in several sections that produce neutrality (all values set to 1).

[1]https://legacy.pandorabots.com/botmaster/en/home
[2]https://www.rivescript.com/

Original Template
<Sin duda |Naturalmente |Normal |Efectivamente |Buena idea |Buena decisión |Fantástico |Genial |Maravilloso |No esta mal |En fin |Claro |Que envidia |En serio? >, <una |cualquiera ><se aburre |no puede estar >tantas horas <|sin salir |sin salir de casa |en casa |sin hacer nada |parada ><|que se agobia |que se cansa >.

Translated Template
<Definitely |Naturally |Normal |Effectively |Good idea |Good decision |Fantastic |Great |Wonderful |Not bad |Anyway |Of course |I envy you |Seriously? >, <anyone |whatever ><gets bored |can not be >so many hours <|without leaving |without leaving home |at home |without doing anything |stand ><|because it is overwhelming |because he gets tired >.

To carry out the elaboration of the dialogue, we begin by identifying the relevant emotions of the current state of the character. Usually, one emotion tends to stand out, but in the case of negative emotions, they tend to stay together. The next step is to analyze the emotional value of each of the combinations of the sentences of our template. From this repertoire of alternatives, we select those in which the same emotions predominate as in the character. Among these, we look for the closest value to the dominating emotion of the character, and then the other emotions following their order.

As stated previously, the quality and precision of our generator depends directly on the quality of both the emotional dictionary and our collection of dialogue templates.

Impact and influence of emotions

The theory of classification states that emotions arise from the interaction between the things we consider important and the events that happen in our environment, which stimulate us (Marsella et al. 2010).

That is why human beings gather events that vary the values of our emotions. It should be noted that a simple conversation in the elevator is considered an interaction and may affect to a greater or lesser degree our emotional state, depending on the content and emotions of the dialogue.

In addition, our emotional state influences the way we talk. Based on the theory of classification, when someone speaks to us, the way they do it and the terms that they use affect our emotions and our response. In turn, the way that person speaks to us is conditioned by their own emotions.

Therefore, we can conclude that, if we speak according to our emotions, and if the way people speak to us affects our response, then the way we speak depends directly on our current emotional state and the emotional state we perceive of the speaker.

Following this assumption, our system must not only take into account the emotions of the character that will speak when generating their response in the dialogue. To be more realistic, we must take into account the response received earlier in the dialogue, with the objective of updating the character's current emotions to select and respond according to the new value of these emotions.

In order to determine how external emotions influence our own emotions, we need to understand how emotions affect one another. For example, sadness is usually affected by emotions such as anger or fear (Bolles and Fanselow 1980). In general, emotions with greater intensity tend to predominate over the others. The relationship between two or more emotions that can suppress the intensity of other emotions is known as inhibition. Some models use techniques that inhibit emotions that are weaker and contrary to those that have greater intensity (Velsquez 1997). For example, sadness would tend to inhibit joy if it had a greater intensity.

To develop our system, we have used a similar technique: opposite emotions are inhibited according to their intensity, but negative emotions (fear, anger, sadness, disgust) have preference over positive emotions (joy) because they are more persistent and dominating emotions.

Therefore, the first thing we do before updating the emotions of a character is to identify if the dominating emotion of the current state of the character is joy or a negative emotion. Similarly, we extract the dominating emotion from the emotions we perceive from the response. If we are cheerful and have a happy interaction, the value of this emotion will increase more than if we have it while we are sad or furious. Likewise, the rest of the emotions will lose more or less weight when getting the answer to the interaction according to our current state. To illustrate these examples, we show the formulas we have obtained after adjusting the thresholds empirically.

If Joy me and NOT Joy sentence:

$$Rel_{new} = Rel_{me} + Rel_{sentence} \cdot 0.4$$

Relevant emotions of the sentence get a small increment, given that my state is contrary to it

$$Joy_{new} = Joy_{me} \cdot 0.8$$

As a result of being an opposite emotion, the negative ones are reduced by 20% as well

If NOT Joy me and NOT Joy sentence:

$$Rel_{new} = Rel_{me} + Rel_{sentence} \cdot 0.6$$

Relevant emotions of the sentence get a big increment, given that my state is complementary to it

$$Joy_{new} = Joy_{me} \cdot 0.5$$

Everything is not-joy, so joy is worth just one half

If Joy me and Joy sentence:

$$Joy_{new} = Joy_{me} + Joy_{sentence} \cdot 0.6$$

$$Other_{new} = Other_{me} \cdot 0.5$$

The remaining emotions are reduced by one half

If NOT Joy me and Joy sentence:

$$Joy_{new} = Joy_{me} + Joy_{sentence} \cdot 0.4$$

$$Other_{new} = Other_{me} \cdot 0.8$$

Getting a joy sentence, the negative ones are reduced by 20%

The stronger emotions tend to dominate the conversation, directly influencing the emotions of the listener. This causes that, during the conversation, the emotions of both characters tend to balance.

Emotions have a temporary life that depends on the weight and importance of the event. In our case, being close events greatly affects our emotions, although being banal conversations, these emotions will take on normality over time. Since our work focuses on short elevator conversations, and how emotions affect each other, we do not evaluate the memory or durability of these emotions. We only focus on how emotions alter the current conversation.

We can observe this in the following tables. In **Table 2** we see a template dialogue where the first character, A, has the emotions [joy: 3.6, anger: 1.4, sadness: 1.1, fear: 1.8, disgust: 1.3], while the character B has the emotions [joy: 1.2, anger: 1.4, sadness: 3.2, fear: 3.0, disgust: 1.3]. **Table 3** shows the result obtained from the same template swapping the emotions of the characters, where it is easy to see that the result of the dialogue is different. In addition, the reader can also see how the emotions of the character influence the response to the previous sentence.

Discussion

As a result of this work, we can emphasize that, although we are using a dictionary obtained from native speakers, there are words with confusing emotions, just as there is the absence of others that we consider emotional, like the days of the week. As long as this dictionary continues to evolve, the results provided by our system are likely to improve.

Among the points that we have not tackeld is the analysis of sentences with negations or sarcastic meanings. These types of sentences need a different approach, since the identification of these factor require additional Natural Language Processing techniques. As with the use of sarcasm, there is a possibility that a character may feel envy. In these cases, when a person feels envy and tries to hide it, he resorts to lies, where the emotions that are perceived in the sentences are the opposite of those he really feels (false happiness).

In order to address all this, it is necessary to take into account the personalities of the characters, which not only helps us recognize selfish characters prone to envy, but there are also types of personalities where the emotional impact of a conversation is different: for example, an empathetic character tends to be more influenced in a conversation than an apathetic one.

Conclusion

The system we have presented is capable of generating short dialogues between two characters, with the content of an elevator conversation. These dialogues present, as their main characteristic, the influence of the emotions of the characters on the dialogue obtained in the end.

During the progress of the conversation, the emotions that the characters perceive from the sentences they receive directly affect their emotional state. This causes not only that the conversations take into account the isolated emotion of each of the characters, but the influence and domination that one type of emotions has over the others. This influence is due to the values of the emotions and their nature, since, as we have previously seen, not all the emotions have the same influence on each other.

Thereby, and taking into account the main objective of this article, which was to create elevator conversations based on the emotions of the characters and the influence they have on that conversation, we can state that the results are satisfactory. We managed to implement a system capable of generating dialogues where emotions are the drivers of the conversation, although we have identified that there is much room for improvement on colloquial expressions, sarcasm and personality of the characters.

Finally, from the two examples mentioned above, we can see in **Tables 4 and 5** two more examples of dialogues created with emotions.

Future Work

As we have seen in the Discusion section, there are many open lines of work for improvement. To highlight some, we believe it is important to address a better phrase analyzer where negation and colloquial expressions are addressed.

Another aspect that must be tackled is to include personality to the characters. In this way we could incorporate factors such as lying, sarcasm, empathy (emotional influence), interest in the other or self-centeredness.

As for the technical improvements, we believe that the next step is to incorporate other NLG techniques that allow us to overcome the limitations of the template-based systems. These systems are very good for rapid prototyping, but they are very expensive to maintain and not very scalable. We will analyze systems based on linguistic dictionaries and machine learning systems based on neural networks.

Finally, we aim to integrate our Emotional Dialogue Generation system as a component that can be used by other storytelling systems such as INES (Concepción, Gervás, and Méndez 2018), with the aim of nourishing and improving narrative generation by incorporating character dialogues.

Character	Generated in Spanish *(Translated into English)*	joy	anger	sadness	fear	disgust
A	Buenas tardes, vecina. Como esta hoy? *(Good afternoon, neighbor. How are you today?)*	3.37	2.33	1	1	1.9
B	Me duele la espalda, pero he salido al parque por la tarde. *(My back hurts, but I went to the park in the afternoon.)*	2.75	1	3.78	1.85	1.4
A	Naturalmente, una se aburre tantas horas aqui. *(It's normal, anyone gets bored for being so many hours here.)*	3.76	1	3.15	1	1
B	Que fastidio, ahora empieza a mejorar el clima. *(How annoying, now the weather begins to improve.)*	4.26	3.25	1	1.15	1.3
A	Disfruta! *(Enjoy!)*	4.45	1	1	1	1

Table 2: A=[joy: 3.6] - B=[sadness: 3.2, fear: 3.0]

Character	Generated in Spanish (Translated into English)	joy	anger	sadness	fear	disgust
A	Me duele la espalda, pero me he bajado al parque por el día. *(My back hurts, but I've gone down to the park for the day.)*	3.80	1	3.83	1.85	1
B	En fin, cualquiera se aburre tantas horas sin hacer nada. *(Anyway, anyone gets bored for being so many hours doing nothing.)*	3.18	1.73	3.92	1.15	1.90
A	Por suerte, ahora empieza a refrescar un poco y creo que se fastidiará la tarde. *(Luckily, now it starts to cool a bit and I think it will spoil the afternoon.)*	3.73	3.25	2.55	1.15	1.3
B	Adios. *(Bye)*	1	1	1	1	1

Table 3: A=[sadness: 3.2, fear: 3.0] - B=[joy: 3.6]

References

Bollen, J.; Mao, H.; and Pepe, A. 2011. Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. *Icwsm* 11:450–453.

Bolles, R. C., and Fanselow, M. S. 1980. A perceptual-defensive-recuperative model of fear and pain. *Behavioral and Brain Sciences* 3(2):291–301.

Cavazza, M., and Charles, F. 2005. Dialogue Generation in Character-based Interactive Storytelling. In *Artificial Intelligence and Interactive Digital Entertainment*, 21–26.

Concepción, E.; Gervás, P.; and Méndez, G. 2018. Ines: A reconstruction of the charade storytelling system using the afanasyev framework. In *Ninth International Conference on Computational Creativity, ICCC 2018*.

Das, D., and Bandyopadhyay, S. 2009. Word to sentence level emotion tagging for bengali blogs. In *Proceedings of the ACL-IJCNLP 2009 Conference Short Papers*, 149152. nullUSA: Association for Computational Linguistics.

Delatorre, P.; Leon, C.; Gervas, P.; and Palomo-Duarte, M. 2017. A computational model of the cognitive impact of decorative elements on the perception of suspense. *Connection Science* 29(4):295–331.

Eugercios, G.; Gutierrez, P.; and Kaloyanova, E. 2018. Anlisis emocional para la inclusin digital. Bsc. thesis, Universidad Complutense de Madrid.

Ferré, P.; Guasch, M.; Martínez-García, N.; Fraga, I.; and Hinojosa, J. A. 2017. Moved by words: Affective ratings for a set of 2,266 spanish words in five discrete emotion categories. *Behavior research methods* 49(3):1082–1094.

Ghosh, S.; Chollet, M.; Laksana, E.; Morency, L.-P.; and Scherer, S. 2017. Affect-lm: A neural language model for customizable affective text generation. *arXiv preprint arXiv:1704.06851*.

Hinojosa, J. A.; Martínez-García, N.; Villalba-García, C.; Fernández-Folgueiras, U.; Sánchez-Carmona, A.; Pozo, M. A.; and Montoro, P. R. 2016. Affective norms of 875 spanish words for five discrete emotional categories and two emotional dimensions. *Behavior research methods* 48(1):272–284.

Lerner, J. S.; Li, Y.; Valdesolo, P.; and Kassam, K. S. 2015. Emotion and decision making. *Annual Review of Psychology* 66(1):799–823. PMID: 25251484.

Mahamood, S., and Reiter, E. 2011. Generating affective natural language for parents of neonatal infants. In *Proc. of*

Character	Generated in Spanish (Translated into English)	joy	anger	sadness	fear	disgust
A	¿Que tal vecino, como va todo? *(What's up neighbor, how is everything going on?)*	2.58	1.65	1.85	1.86	1.91
B	Buenos días amigo. Pues mi madre viene por vacaciones *(Good morning friend. Well, my mother is coming for a vacation)*	3.82	1	1	1	1
A	Fantástico, me alegro mucho, disfruta los buenos momentos *(Fantastic, I'm glad, enjoy the good times)*	4.73	1	1	1	1
B	Se agradece, un saludo *(Thank you, best regards)*	4.39	1	1	1	1

Table 4: A=[joy: 2.6] - B=[joy: 2.2, disgust: 2.3]

Character	Generated in Spanish (Translated into English)	joy	anger	sadness	fear	disgust
A	Que tal, ¿sabes cuando empieza el partido hoy? *(How are you, do you know when the game starts today?)*	4.76	1.43	1.3525	1.47	1.37
B	Buenas, empieza a las cinco y media. *(Hi, it starts at five thirty.)*	3.35	1	1.6	1	1
A	Genial me iré un poco más pronto con la idea de intentar llegar a tiempo. *(Great I'll leave a little sooner with the idea of trying to arrive on time.)*	3.173	1	2.55	1	1
B	Te dará tiempo. *(You will be on time.)*	3.3	1	1	1	1

Table 5: A=[joy: 4.2] - B=[sadness: 3.6]

the 13th European Workshop on Natural Language Generation, 12–21. Association for Computational Linguistics.

Marsella, S.; Gratch, J.; Petta, P.; et al. 2010. Computational models of emotion. *A Blueprint for Affective Computing-A sourcebook and manual* 11(1):21–46.

Méndez, G.; Gervás, P.; and León, C. 2016. *On the Use of Character Affinities for Story Plot Generation*, volume 416 of *Advances in Intelligent Systems and Computing*. Springer. chapter 15, 211–225.

Minsky, M. 2010. La máquina de las emociones. *Del dolor al sufrimiento* 3:92–125.

Miranda, C. N. H.; Luna, J. A. G.; and Morillo, D. S. 2016. Minería de opiniones basado en la adaptación al español de anew sobre opiniones acerca de hoteles. *Procesamiento del Lenguaje Natural* 56:25–32.

Oh, A. H., and Rudnicky, A. I. 2000. Stochastic language generation for spoken dialogue systems. In *Proc. 2000 ANLP/NAACL Workshop on Conversational systems-Volume 3*, 27–32. Assoc. for Computational Linguistics.

Oñate, A.; Méndez, G.; and Gervás, P. 2019. Introducing mood and affinity to generate brief template-based dialogues in storytelling systems. In *C3GI: The 7th International Workshop on Computational Creativity, Concept Invention, and General Intelligence*. Bozen-Bolzano, Italy: CEUR Workshop Proceedings.

Pak, A., and Paroubek, P. 2010. Twitter as a corpus for sentiment analysis and opinion mining. In *LREc*, 1320–1326.

Polzin, T. S., and Waibel, A. 2000. Emotion-sensitive human-computer interfaces. In *ISCA tutorial and research workshop (ITRW) on speech and emotion*.

Quan, C., and Ren, F. 2010. Sentence emotion analysis and recognition based on emotion words using ren-cecps. *International Journal of Advanced Intelligence* 2(1):105–117.

Shawar, B. A., and Atwell, E. 2007. Different measurements metrics to evaluate a chatbot system. In *Proceedings of the workshop on bridging the gap: Academic and industrial research in dialog technologies*, 89–96. Association for Computational Linguistics.

Stadthagen-González, H.; Ferré, P.; Pérez-Sánchez, M. A.; Imbault, C.; and Hinojosa, J. A. 2018. Norms for 10,491 spanish words for five discrete emotions: Happiness, disgust, anger, fear, and sadness. *Behavior research methods* 50(5):1943–1952.

Traum, D. 2017. Computational approaches to dialogue. *The Routledge Handbook of Language and Dialogue* 143.

Velsquez, J. 1997. Modeling emotions and other motivations in synthetic agents. *Aaai/iaai* 10–15.

Wang, W.; Chen, L.; Thirunarayan, K.; and Sheth, A. P. 2012. Harnessing twitter" big data" for automatic emotion identification. In *Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on and 2012 International Confernece on Social Computing (SocialCom)*, 587–592. IEEE.

Zhou, H.; Huang, M.; Zhang, T.; Zhu, X.; and Liu, B. 2018. Emotional chatting machine: Emotional conversation generation with internal and external memory. In *Thirty-Second AAAI Conference on Artificial Intelligence*.