A Procedurally Generated Approach to Emotional Storytelling for Games and Interactive Systems

Duarte Dinis Dordio Parreira Ferreira Cabral

Instituto Superior Técnico - Campus Taguspark, 2744-016 Porto Salvo, Portugal
Advisor: Professor Carlos António Roque Martinho

Abstract. One of the least noticeable applications in the area of Procedural Generated Content is in generating narratives. Unfortunately, this results in a disregard for the potential emotional impact that storytelling may bring to table, since that, in many cases, it is the way a narrative is told that results in some sort of emotional resonance for an audience. Through a study of storytelling techniques, together with an analysis of existing narrative-driven PCG systems, and by exploring various manners of emotional evaluation in art, a system was built that takes base narratives as input and, through an emotional model, adds storytelling mechanisms to them, seeking to differentiate and maximize one of two opposite emotional moods: Happy and Dour. These narratives are then played in the Virtual Tutor system. A series of experiments consisting of four different narratives were made to discern if the emotional valence of these was successfully differentiated and maximized.

Keywords: Affective Computing · Emotions · Storytelling · Narrative · Turn-taking.

1 Introduction

An area in game design where PCG is less knowingly used is in story generation, and that is mostly because it usually is subservient to other forms of procedural generated content. Usually here, what is actually generated is a *plot*, the sequential events within a certain narrative domain that are originated through some combination of character related decisions. As Reidl and Young claim [17], the reason plot is targeted in this area is because it essentially derives from a logical causal progression and believable character motivations, which is very appropriate given the way PCG systems tend to work.

Due to this however, *storytelling*, the way the narrative is in fact told and expressed towards the user of the system, may end up being, for the most part, just a straightforward and direct recounting of events. A narrative with an well-written plot but a too simplistic or inappropriate approach to storytelling may end up feeling emotionally unengaging to an audience.

With this work, I aim to create a narrative manipulation model that generates the best possible version of a given narrative according to a binary emotional spectrum or mood (Happy or Dour).

Through this, my hope is to bring to light the importance of having storytelling as a main focus in procedurally generated narratives, by showing that the right storytelling approach to a given narrative may result in more palpable and intensely felt emotions;

2 Related Work

2.1 Building Blocks of a Narrative

In Figure 1, a visual representation of what classically constitutes a narrative can be seen. Starting from the most basic concept, a *story* is the properly ordered timeline of relevant events inserted in a narrative. In his book, Forster described that, for the reader, the process of reading a simple story is based on the constant asking of "and then **what** happened?" [5].

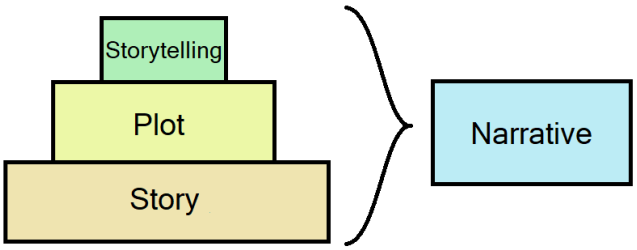

Fig. 1. The narrative pyramid

For its second level, we have the *plot* of the narrative, which takes the foundation lain by the story and adds a sense of causality to it. Meaning, it details the reasons and consequences for why the major events of the story happened in the first place. Forster describes that, for a plot, the reader finds him/herself in a continuous cycle of asking "**how** and **why** did that happen?" [5]. Numerous different plots may originate thus from a single story.

The topmost level of the pyramid, *storytelling*, is the chosen manner in the plot is relayed to the audience. For each successive plot point, it can be viewed as the an answer to "how am I told this?" With these three ideas combined, a proper *narrative* is defined.

2.2 Interactive Stories

Nowadays, more and more focus is being given to development of systems that properly allow users to experience interactive narratives, which are normally

called **Interactive Stories/Dramas**. These types of narratives tend to differentiate themselves from other more traditional types by having an explicit focus in presenting a multitude of possible ways in which they may develop, branch out and conclude themselves. A few of the main examples of Interactive Stories of the past 15 years are *Façade* [13], *Prom Week* [14] and *Bad News* [15].

2.3 Models for Procedurally Generated Storytelling

Two planners that indirectly informed my work for this project were the *Suspenser* [4] and *Prevoyant* systems [1].

Cheong and Young's *Suspenser* system seeks to develop narratives that are specifically built to arouse suspense on the reader, through a combination of a planning algorithm with psychological theories about human mental cognition. By selecting which story elements are presented, the imaginary state and outcome of the story in the mind of the reader will be altered, resulting on feelings of "excitement or anxiety". For a given instance of a *fabula* (plot of a narrative), the system is a generator of potential *sjuzhets* (manners in which the *fabula* is transmitted to a viewer, i.e. storytelling), with a proper order and form.

Prevoyant (an extension of the Suspenser model) that introduced two storytelling elements to a base plot: Flashbacks and foreshadowing. The main motivation for this work was that "emotions such as suspense, curiosity, and surprise help the readers focus attention on a story, contributing to the readers' sense of satisfaction". As it is mentioned in the referenced article, "authors can let the readers know about some facts in advance or hide some information until a certain point for a dramatic effect".

Both of these models offer a proper overview on how to implement storytelling aspects to an existing plot, which, while extremely relevant for my work, are too overly complicated for me to try to emulate them in my work. Instead, they served more as to conceptually guide some parts of my proposed solution for this project.

2.4 Studies of Emotions in Interactive Stories

These types of studies are advantageous in two correlated ways: the first, and most obvious, is it allows for a better understanding of how the human mind entangles itself into a narrative space and begins reacting in relation to its evolution. But if you look carefully between the "cracks" of this virtue, you can see another one that I would argue is just as important and begs for further attention: it allows us to tell better, more involving and rich stories, ones that focus their efforts in specific topics and ideas so as to have their intended audience always engaged.

Studies such as the one made by Jakobsen, Christensen and Bruni explore emotional empathy through the choice and moral alignment [9], while Girina [6] showed the many ways the cinematic language influences emotions through concepts such as expressive lighting and proper staging.

2.5 Emotional Evaluation and Measurement for Art

A few of the evaluation techniques for artistic media (film, literature, video games, music) I took into consideration for my own project are mentioned next, such as Hevner's Adjective Circle [7][21], NFA [12], and SAM [10]. The most relevant one for the work, however, would be the one used by Schaefer, Nils, Sanchez and Philippot [19], where film clips were emotionally evaluated by three groups of questions:

- **Self-reported emotional arousal** - A self-assessed measurement of the emotional arousal the subject suffered during the movie clip.
- **Differential Emotions Scale** - To measure discrete emotional dimensions, a modified version of Differential Emotions Scale (DES) [8] was used, with the original 10 groups of emotional adjectives plus 6 additional ones. 5 of these groups consisted of positive emotions, 8 of negative ones, and the remaining 3 consist of emotions with high arousal, but no specific valence.
- **PANAS** - This scale is used as a self-reported adjective checklist [22] containing two separate 10-item subscales of emotions, one for Positive Affect (PA) and other for Negative Affect (NA).

This was the evaluation method mainly used in the experimental context of this project.

3 Implementation

3.1 Development Tools

Yarn - For the creation of the narratives my system works with, a dialogue editor called *Yarn*[1] was used. It allows its user to easily create branching instances of dialogue between multiple characters and the user. Also, through directed connections, a variety of scripted dialogues paths can occur.

The file containing the contents of a series of yarn dialogues is called the **yarnfile**. Its contents are interpreted by the Yarn program and it generates the appropriate visual nodes and connections. In this work however, manipulation of the yarnfile happens directly on the file itself. A node is made up of a title, a set of tags and a body of text.

Virtual Tutor - VT is an Android app developed by 4 students of Instituto Superior Técnico, with 3 of them working on this tool in the context of their own projects for the master's degree [11] [16] [20].

This app was built with the purpose of helping students organize their study, while also providing feedback regarding their academic results that could be useful [18]. This is done by two virtual characters who serve as the personal tutors of the user. For the context of this work, however, these two characters are referenced as **NPC Storytellers**, or simply as narrators. These are named Maria and João.

[1] I. A. Inc. (2016) Yarn.

Unity - Both of the systems mentioned above are conjointly used with the Unity Engine[2] (Version 2018.3.2).

3.2 Storytelling System

My storytelling system serves as added extension to the VT system. The Storytelling process is divided into four distinct main phases, all sequenced and organized through the Storytelling Manager of the system.

Yarnfile Processing - The user selects the path for the desired yarnfile (one that a user created before using the system), the desired emotional mood (Happy or Dour), or the desired Narration Style for the NPC Storytellers.

The system reads through, extracts and stores the narrative body and tags of the yarnfile. One relevant set of tags is used inside the narrative body of each node, but these will be detailed further on, since they relate with the developed emotional system.

The end result of this phase is a list of nodes that will be used by the following phase of the Storytelling pipeline.

Narrative Generation - After obtaining the list of nodes, comes the process of obtaining the **best version** for the narrative. If the user wants a narrative with a "Happy" mood, the nodes of the narrative considered by the Emotional Model as being more depressing/violent/sad are less likely to appear in the chosen "best version" of the narrative. The opposite idea applies if the player chooses a narrative with a "Dour" mood. Note that a node may only be omitted if one of its tags describes it as being a non-essential node.

A more thorough description of what happens on this stage can be seen in Subsection 3.3.

Post-Generation Fixes - After having selected the most suited narrative (i.e. set of nodes) for the user-selected emotional mood, this phase of the pipeline adds a couple of fixes to the yarnfile.

The first changes the way the two NPC Storytellers will switch the narration of the narrative between themselves in VT (in a periodic fashion, or according with the evolution of the mood along the narrative). Next, emotional expressions tags are added to the nodes in order to have the NPC Storytellers visually emote according with the emotional developments of the narrative.

These fixes will be explained in Subsection 3.3.

Yarnfile Generation This final phase simply consists in exporting the set of nodes into a new yarnfile and signaling the VT system that it can begin processing it. After this, the user can run the generated narrative in Virtual Tutor.

[2] U. Engine. (2005) Unity technologies.

3.3 Emotional Model

The core element of the system I developed lies in the Emotional evaluation system. This model essentially allows the system to infer the overall emotional mood of a narrative. In the course of this Section, all of the components that go into this will be fully explained.

Objective Tags - We use objective(s) of the main character(s) in a narrative as the basis for the emotional system. If a character is getting close to its objectives, then it can be infered that there is a happier mood to that part of the narrative. If a character is completely out-of-reach of any of its objectives, then the mood of the narrative is interpreted as being darker and more dour.

In the yarnfile, this is represented by adding special tags with the format of **<<OBJECTIVE:XX, sec_objective:YY, ...>>** in the Body Section of each node whenever changes in the objectives of a character occur. An objective is represented by an upper-cased name (**OBJECTIVE**) if it is a main objective and by a non-spaced, lower-cased custom name for the secondary objectives (**sec_objective**), both of these types having a value ranging from -1 to 1 saying just how close to that objective the main character is (**XX, YY**). The manner in which the objectives of a character vary throughout the narrative has to be done manually done by the writer of the narrative.

Choosing the Best Narrative - For the first step, what the Emotional Model does is essentially a joint evaluation of the values of every single consecutive group of 3 objective tags throughout the entire course of the narrative. An Objective Value is calculated by adding the full value of each main objective with half the value of each secondary objective. From here onward, the first Objective Value will be called $obj1$, the second $obj2$ and the third $obj3$. This is a process that is executed from the very first existing objective tag in the narrative, meaning that in the first iteration of the process there will be two cases Objective Values have to be pre-defined in order to make up for their non-existence. On the second iteration, there will still exist one Objective Value that needs pre-defining. So, in these cases, these Objective Values are assigned with 0.

For the second step, the Model classifies the manner in which the trio of Objective Tags vary with a corresponding Mood Variation value. As seen in Figure 2, there are 9 possible types of variation in values, each with an associated integer value ranging from -3 to 3. The Mood Variation value represents the valence for the mood in a group of three objective tags. A negative value means negative valence (more Dour), and a positive one means positive valence (more Happy). A special mention to the images referring to the values of -3 and 3 valence. Those are considered to be negative or positive surprises (respectively) that may occur in the narrative, and that is why they are valued above below and above all other Mood Variation values.

After obtaining the Mood Variation value, we generate the Current Mood value. This is done by calculating the absolute value of the difference between

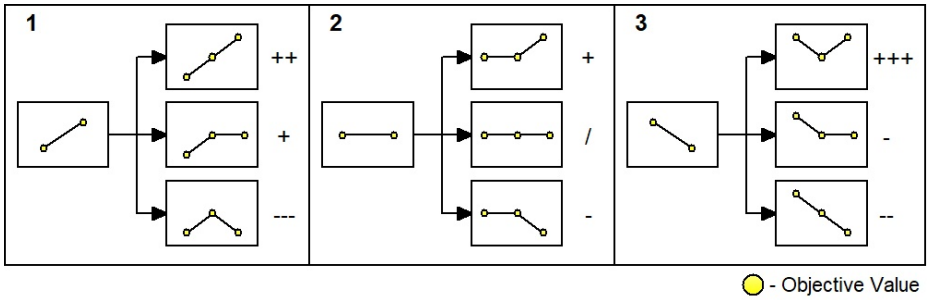

Fig. 2. Diagram of the various types of Mood Variation, along with the respective classification (/ is 0).

$obj3$ and $obj1$, and then multiplying it by the Mood Variation value of the current group of 3 Objective Values. The formula for this is the following:

$$CurrentMoodValue = (obj3 - obj1) * MoodVariationValue$$

This is a required calculation because we need a precise manner in which to measure the exact way the trio of Objective Values varies. When all groups of three consecutive Objective Values are evaluated, then the final TEV is obtained and the narrative evaluation ends. A Happy narrative is expected to have a high positive TEV, and a Dour narrative is expected to have a low negative TEV.

This is how the Emotional Model operates in the context of the second phase of the Storytelling Pipeline of Subsection 3.2.

Narration Exchanges and Emotional Expressions - This emotional model generates a list that stores the ways in which the pairs of Mood Variation and Current Mood values evolve along the selected best narrative. These values are used in the **Post-Generation Fixes** phase of the pipeline.

The first of these fixes changes the way in which the two NPC Storytellers alternate between themselves while narrating the narrative, which may enhance the emotional impact of the narrative if done in the right way. Depending on what the user of the system chooses, the NPC Storytellers may alternate between themselves from node to node or switch according to the variations in the mood of the narrative. Focusing specifically on the latter style, by using the list of paired values mentioned in the beginning of this Subsection, more specifically the way in which the Current Mood values evolve along the narrative, the system alternates the narrator in the yarnfile according with the manner in which the mood evolves in the narrative.

When a noticeable increase happens in two successive Current Mood values and the current narrator is João, Maria becomes the new narrator. If instead of this a noticeable decrease in the two successive Current Mood values happens and Maria is the current narrator, then the current narrator is switched to João. Otherwise, the current narrator is maintained.

The second type of fix/add-on that uses the list of paired values is the one that introduces command tags in the narrative body of each node for the NPC Storytellers to visually emote according with the emotional mood the narrative is taking.

The NPC Storytellers of the VT system support a total of 7 emotions with values ranging between 0 and 1, but for the context of this narrative we unly use 5: Neutral, Happiness, Sadness, Fear and Surprise (Anger and Disgust were discarded for the context of this work, seeing that they did not fit inside the context of the narrative used during the testing process). Maria was assigned the possible emotions of Happiness and Sadness (focus on valence), while João could emote Fear and Surprise (focus on arousal). Note that both NPC Storytellers are allowed to express the Neutral emotion.

These emotions are expressed by a combination of simulated facial expressions (developed by Silva [20]) along with stylized dialogue balloons (developed by Pereira [16]). Each of the listed Mood Variation values of the narrative is used whenever its respective Objective Value tag is reached, so the system can properly introduce emotional expressions in those precise moments of the narrative body.

Mood Variation Value	Maria	João
-3	Sadness: 1.0	Fear: 1.0
-2	Sadness: 0.5	Fear: 0.5
-1	Sadness: 0.2	Neutral: 0.5
0	Neutral: 1.0	Neutral: 1.0
1	Happy: 0.2	Neutral: 0.5
2	Happy: 0.5	Surprise: 0.5
3	Happy: 1.0	Surprise: 1.0

Table 1. Table describing how the emotions displayed by the NPC Storytellers change according to Mood Variation values.

In Table 1, the correspondence between each Mood Variation value and the respective emotional expression assignments is specified.

It is through this combination of emotionally focused systems that this work seeks to prove the effectiveness of storytelling in the emotional impact of a narrative. The choices of switching narrator, having them visually emote and react to specific occurrences, finding the best intended version of the narrative with the right amount of sections omitted, all of it is targeted at enhancing the narrative experiences of an audience.

4 Experimental Evaluation

4.1 Preliminary Evaluation

At an earlier stage of the development of this project, a preliminary experience was conducted in order to check whether a distinction in emotions could be verified in the Happy and Dour versions of an input narrative (this one being a longer and more detailed version of the one used in the final experiment).

For this, a small group of 5 subjects watched in succession two videos of the VT system running the two interpretations of the narrative. After this, they answered a small questionnaire with a list of 15 adjectives inspired in the Hevner's Adjective Circle [7] to determine if there was any immediate difference between the two narratives.

The results were favorable for the continuation of the experiment, with a suggestive difference in the values of positive adjectives being seen across the Happy and Dour narratives.

4.2 Final Experiment: Materials

Narratives - 4 different sets of tests and questionnaires were devised with 4 different narratives contained within them. These narratives are a text only version of the narrative with a **Happy** mood (**T+**); a text only version of the narrative with a **Dour** mood (**T-**); a video of the narrative with a **Dour** mood, where the current NPC Storyteller narrator is selected according with to **progression in mood** inside the narrative (**E-**); and a video of the narrative with a **Dour** mood, where the NPC Storytellers switch narration on **node to node** basis (**A-**);

The base narrative plot that was used as an input was written by myself.

Questionnaires - Each of the 4 different narratives has an associated questionnaire, all of them with identical questions.

After an initial greeting and asking the age and gender of the suer, test-subjects are prompted with a set of 8 questions (4 positive questions, and 4 negative ones) taken from the NEO template from IPIP that evaluates the "Openness to Experience"[3] from a person (Likert Scale of 1 to 7, 1 - "Completely disagree" and 7 - "Completely agree").

Then, the textual/video narrative is presented to the subject. After having experienced it, a pair of questions regarding the perceived emotional arousal are presented. The first was directly brought from the work of Schaefer et al. [19] (Likert Scale of 1 to 7, 1 - "I felt no emotions at all" and 7 - "I felt very intense emotions"), and the other used the arousal segment of the SAM questionnaire [10] seen in the bottom half of Figure **??** (Likert Scale of 1 to 7, 1 - "I felt no emotions at all" and 7 - "I felt very intense emotions").

After this, 16 different groups of adjectives from the DES questionnaire [19] are presented, with each group being similar in meaning and feeling. These vary

[3] https://ipip.ori.org/newNEOKey.htm

in terms of valence and arousal (Likert Scale of 1 to 7, 1 - "Not at all" and 7 - "Very intensely").

Then, 20 different emotions (half being positive, the other half negative) are presented from the PANAS questionnaire [19]. The user must describe how these matched the emotions they felt during the course of the narrative (Likert scale of 1 to 5, 1 - "Very slightly or not at all" and 5 - "Extremely").

Finally, a final section verifies whether the user was in some way not paying attention to the narrative in the questionnaire by posing a trio of questions regarding details of the narrative.

4.3 Final Experiment: Results

Note first that the results from the questionnaires were analysed in four major ways. First, **T+** and **T-** were used to compare the differences in emotional valence for two narratives with opposite emotional moods; **E-** and **A-** are compared to check whether the added nuances to the manner in which the NPC Storytellers alternate between themselves while narrating; and finally, **T-** is compared to both **E-** and **A-** to see which of the two storytelling forms was more emotionally effective (text vs video).

Demographic Analysis - A total amount of 48 participants (12 per questionnaire) made up the entirety of the Final Experience. All of the gathered data from the 4 questionnaires was analysed and processed through SPSS Statistics (version 25). From the total of 48 participants, a total of 35 classified their gender as Male (72.9%), and 12 said they were of Female gender (25%). The gender of the last remaining participant was agender, which was classified as Other (2.1%).

The Openness of each subject was calculated with the following formula: $Openness(q) = pos1(q) + pos2(q) + pos3(q) + pos4(q) - pos5(q) - pos6(q) - pos7(q) - pos8(q) + 32$, which results in a minimum value of 8 and a maximum value of 56. The results showed a mean value $\overline{x} = 46.33$ and standard deviation $s = 5.471$ (subjects with a general openness to artistic media).

Arousal - A T-Student test was used to verify if there were any significant differences between the total arousal values of the four narrative pairings (T-Student was chosen because the samples followed normal distributions).

The only relevant result from this T-Student test was when comparing the **T-** and **E-** narratives, where the result was ($T(22) = 2.902$, $p = 0.008$). From which we can conclude that the average total arousal for the **T-** narrative is significantly higher than the one for **E-**.

PANAS - An average value was calculated for both halves of the PANAS variable (paAvg and naAvg).

Again, a T-Student test was used to check whether the values of paAvg and naAvg differed significantly between the four comparative pairings of narratives (because the Shapiro-Wilk test between concluded that each had a normal distribution, i.e. $p < 0.05$). From this, it was noted that a significant difference in

the values of paAvg happened between **T-** and **E-** ($T(22) = 2.593$, $p = 0.017$), and also between **T-** and **A-** ($T(22) = 2.538$, $p = 0.019$). More specifically, the **T-** narrative had a significantly larger paAvg mean value than its video counterparts.

A non-parametric Spearman's Correlation showed that the values of paAvg and naAvg had a strong correlation with the total arousal value, with the corresponding results for each being ($r_s(48) = 0.601$, $p = 0.000$) and ($r_s(48) = 0.551$, $p = 0.000$).

A 2-tailed Spearman's Correlation test verified that there was also a strong correlation between the values of paAvg and naAvg ($r_s(48) = 0.430$, $p = 0.002$).

Spearman's Correlation test also suggests that the paAvg value of a subject has a slight tendency to increase along with his/her total Openness value ($r_s(48) = 0.263$, $p = 0.036$).

DES - 5 of the 16 adjective groups had positive emotions/adjectives, and 8 groups had negative ones, so the mean values for each of these major groups (desPosAvg and desNegAvg) were used for most of this part of the statistical analysis.

T-Student test was run for the 4 relevant narrative pairings. A significant difference in mean values of desPosAvg was found between the **T-** and **E-** narratives ($T(22) = 2.145$, $p = 0.043$).

It was seen that these have a strong positive correlation with the paAvg and naAvg values from PANAS. All of the four possible comparisons between these values, done with a Spearman's Correlation test (both 1-Tailed and 2-Tailed), resulted in p values equal or inferior to 0.02, which in itself is inferior to 0.05. For desPosAvg, the exact results for the Spearman test with paAvg were ($r_s(48) = 0.706$, $p_{1-tailed} = 0.000$, $p_{2-tailed} = 0.000$), and with naAvg were ($r_s(48) = 0.399$, $p_{1-tailed} = 0.002$, $p_{2-tailed} = 0.005$). For desNegAvg, the corresponding results were of ($r_s(48) = 0.336$, $p_{1-tailed} = 0.02$, $p_{2-tailed} = 0.01$) for paAvg, and ($r_s(48) = 0.795$, $p_{1-tailed} = 0.000$, $p_{2-tailed} = 0.000$) for naAvg.

DES also has a very strong positive correlation with the total values for arousal. For desPosAvg, the exact results of the Spearman's Correlation test with the arousal parameter were ($r_s(48) = 0.505$, $p_{1-tailed} = 0.000$, $p_{2-tailed} = 0.000$), and the respective results for desNegAvg were ($r_s(48) = 0.570$, $p_{1-tailed} = 0.000$, $p_{2-tailed} = 0.000$)

As in similar fashion with PANAS, DES presents a direct positive correlation in its desPosAvg and desNegAvg values, with a Two-tailed Spearman's Correlation test resulting in ($r_s(48) = 0.368$, $p = 0.010$).

The positive mean for DES also, similarly as with paAvg value of PANAS, presents a direct positive correlation with the Openness value, with a Spearman's Correlation test resulting in ($r_s(48) = 0.368$, $p = 0.010$).

Verification Questions - When running a Kruskal-Wallis H test to check whether some significant difference existed between mean values for **T+**, **T-**, **T+** and **A-**. The results were ($\chi^2(3) = 10.073, p = 0.018$), which means that some significant difference exists between these 4 values. A Mann-Whitney U test

across all pairs of results (with an $\alpha = 0.05/4 = 0.0125$) resulted in a significant difference existing between the results of **T+** and **A-**, and also **T-** and **A-**, with the result being ($U = 42$, $p = 0.014$).

4.4 Results

When contrasting the results of the **T+** narrative with the ones from the **T-**, meaning that no visible difference in the moods of the narratives was seen.

There was no significant difference in the results obtained by comparing detected between the two video narratives, **E-** and **A-**.

There was also no significant difference in results between the **T-** and **A-** narratives

The **E-** narrative was the one that, according to the expectations for this work, was thought to generate a stronger intensity in terms of negative (Dour) emotions, meaning greater values regarding negative valence emotions and a higher arousal than the other two narratives. However, the exact opposite happened, with the **E-** narrative being the uncontested "loser" of this series of experiments, with its results being consistently worse compared to **T-**.

5 Conclusions

Using the Virtual Tutor system along with the Yarn tool, I developed a system that received as input properly tagged and labeled narratives that were then processed according with choices the user of the system made, resulting in narratives that tried to maximize one of two emotional moods (Happy or Dour), that sought as well to maximize the intensity of the emotions they brought on a spectator. This system supported itself on an emotional model that was built from the ground up to support the ideas and objectives of this work.

To test this idea, four different versions of the same input narrative were created that could be use to test the hypotheses raised. Unfortunately, the results were inconclusive regarding the effective distinction in emotional moods towards a base narrative, and showcased that narratives run in the VT system had a lesser emotional impact on subjects. But despite this, it was proven that the manner in which a narrative is told can directly influence the intensity of the emotions it is liable to trigger. The manner in which a narrative is told can indeed enhance its emotional results, but could also do the total opposite and diminish its overall impact, both in terms of arousal and valence.

To continue the work done here, I find it would be necessary to pinpoint which of the many parts of the system failed the Emotional Model and improve from there.

References

1. Bae, B. and Young, R. "A use of flashback and foreshadowing for surprise arousal in narrative using a plan-based approach," in Interactive Storytelling, U. Spierling and N. Szilas, Eds. Berlin,Heidelberg: Springer Berlin Heidelberg, 2008, pp. 156–167.

2. Bartsch, A. and Appel, M. and Storch, D. "Predicting emotions and meta-emotions at the movies: Therole of the need for affect in audiences' experience of horror and drama,"Communication Research- COMMUN RES, vol. 37, pp. 167–190, 2010
3. Battaglino, C. and Damiano, R. and Lombardo, V. "Moral values in narrative characters: An experiment in the generation of moral emotions," in Interactive Storytelling, A. Mitchell, C. Fernández-Vara, and D. Thue, Eds. Cham: Springer International Publishing, 2014, pp. 212–215.
4. Cheong, Y. and Young, R. "Narrative generation for suspense: Modeling and evaluation," in Interactive Storytelling, U. Spierling and N. Szilas, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg,2008, pp. 144–155
5. Forster, E. (1927) Aspects of the novel.
6. Girina, I. "Video game mise-en-scene remediation of cinematic codes in video games," inInteractiveStorytelling. Cham: Springer International Publishing, 2013, pp. 45–54
7. Hevner, K. "Experimental studies of the elements of expression in music," vol. 48, 1936, pp. 246–268.
8. Izard, C. and Dougherty, F. and Bloxom, B. and Kotsch, N. "The differential emotions scale: A method ofmeasuring the meaning of subjective experience of discrete emotions," 1974
9. Jakobsen, M. and Christensen, D.S. and Bruni, L.E. "How knowledge of the player character's align-ment affect decision making in an interactive narrative," inInteractive Storytelling, N. Nunes, I. Oak-ley, and V. Nisi, Eds. Cham: Springer International Publishing, 2017, pp. 193–205.
10. Lang, P. "Behavioral treatment and bio-behavioral assessment: Computer applications," pp. 119 –137, 1980.
11. Lima, A. "Virtual tutor: Virtual tutoring agent using empathy and rapport techniques," 2017.
12. Maio, G. and Esses, V. "The need for affect: Individual differences in the motivation to approach oravoid emotions,"Journal of Personality, vol. 69, pp. 684–696, 2001
13. Mateas, M. and Stern, A. (2006) Façade. Playabl Studios LLC.
14. McCoy, J. and Treanor, M. and Samuel, B. and Mateas, M. and Wardrip-Fruin, N. (2012) Prom week. Center for Games and Playable Media.
15. McCoy, J. and Treanor, M. and Samuel, B. and Tearse, B. and Mateas, M. and Wardrip-Fruin, Noah. (2016) Bad news. Expressive Intelligence Studio.
16. Pereira, R. "Expressing emotions through animated speech balloons," 2018
17. Reidl, M. and Young, R. (2010) Narrative planning: Balancing plot and character. Journal of Artificial Intelligence Research 39.
18. Rodrigues, R. and Silva, R. and Pereira, R. and Martinho, C. "Interactive empathic virtual coaches based on the social regulatory cycle," 2019
19. Schaefer, A. and Nils, F. and Sanchez, X. and Philippot, P. "Assessing the effectiveness of a large database of emotion-eliciting films: A new tool for emotion researchers,"Cognition and Emotion, vol. 24,no. 7, pp. 1153–1172, 2010.
20. Silva, R. "Supporting affective expression in multi-party interactions," 2018.
21. Schubert, E. "Update of the Hevner Adjective Checklist,"Perceptual and Motor Skills, vol. 96, no.3suppl, pp. 1117–1122, 2003.
22. Watson, D. and Clark, L. and Tellegen, A. "Development and validation of brief measures of positive and negative affect: The PANAS scales, "Journal of Personality & Social Psychology, vol. 54, no. 6, pp.1063–1070, 1988
23. Young, R. and Pollack, M. and Moore, J. "Decomposition and causality in partial-order planning," 1994