Framing Tension for Game Generation

Phil Lopes, Antonios Liapis, Georgios N. Yannakakis
Institute of Digital Games, University of Malta, Msida, Malta
{louis.p.lopes; antonios.liapis; georgios.yannakakis}@um.edu.mt

Abstract

Emotional progression in narratives is carefully structured by human authors to create unexpected and exciting situations, often culminating in a climactic moment. This paper explores how an autonomous computational designer can create frames of tension which guide the procedural creation of levels and their soundscapes in a digital horror game. Using narrative concepts, the autonomous designer can describe an intended experience that the automated level generator must adhere to. The level generator interprets this intent, bound by the possibilities and constraints of the game. The tension of the generated level guides the allocation of sounds in the level, using a crowdsourced model of tension.

Introduction

Several computationally creative systems have stood at the interplay of different multidisciplinary creative domains. It should not come as a surprise, therefore, that several projects in computational creativity tackle the transformation of data from one domain to another, e.g. images to soundscapes (Johnson and Ventura 2014), news articles to collages (Krzeczkowska et al. 2010), academic papers to songs and their lyrics (Scirea et al. 2015), text descriptions to player abilities (Cook and Colton 2014), to name a few. Due to the dissimilarities between source and target creative domains, such computational systems must learn to creatively interpret the patterns of the input, and work towards making them apparent in the output while still obeying the constraints and the expressivity of the target creative domain (e.g. a limited color palette).

In this context, digital games are particularly relevant as a multi-faceted medium where visuals, audio, narrative and rule- and level-design come together in an interactive experience (Liapis, Yannakakis, and Togelius 2014). Not only must these creative domains go well together, but they must provide players with an enjoyable experience: depending on the genre, this experience can be, for instance, frantic in "bullet hell" action games, relaxing in exploration games, or tense in horror games (Ekman and Lankoski 2009).

When drawing inspiration from dissimilar creative domains, it is important to find the right patterns to replicate (or re-interpret) in the creative output of the system. While systems can look at structural similarities and associations (Grace, Gero, and Saunders 2012), a promising approach is to identify the intentions of the creator of one artefact and attempt to match those intentions in the artefact of the other domain. Towards that outcome, having access to a *frame of reference* for the intentions going into the creative act is ideal. Framing information, as suggested in the FACE model of Colton, Charnley, and Pease (2011), can be provided by the creative system itself as "a piece of natural language text that is comprehensible by people". Such framing information can clarify the intentions of the system in its design choices and can make its creativity more easily perceptible (Colton 2008). Moreover, the framing information can act as a guide when transforming media generated by such a creative system into different media.

In the context of digital games, a human game designer's primary concern and frame of reference is the intended player experience. In most games, the intended player experience affects all design decisions: from the color palette to the responsiveness of the controls and from the sound effects for rewards to the back-story presented in an introductory cut-scene. Taking a successful horror game such as *Amnesia: The Dark Descent* (Frictional Games 2010) as an example, the intended player experience is one of dread, of imminent tragedy, of confusion and constant second-guessing of players' perception and actions. Towards this experience, the visuals include dark colors and dim lights, the audio focuses on ambient noises which foreshadow monsters, the level design has narrow corridors and low visibility while the game rules preclude any way to combat monsters.

This paper extends the *Sonancia* creative system, (Lopes, Liapis, and Yannakakis 2015a; 2015b) by providing the software with the capacity to choose and describe the intended player experience, which is then used to generate game levels and their soundscapes for a horror game. The ability of the computational designer to describe its intentions in clear text to a human audience is paramount in the perception of creativity. Moreover, the system can then create the frame (as the progression of tension) via evolutionary search driven by several fitnesses targeting specific narrative structures. The paper includes several examples of generated frames and their corresponding levels and soundscapes. As an additional contribution to earlier work, the current version of *Sonancia* uses a crowdsourced model of tension to allocate sounds to the level in a way that more closely matches the human perception (or ground truth) of tension.

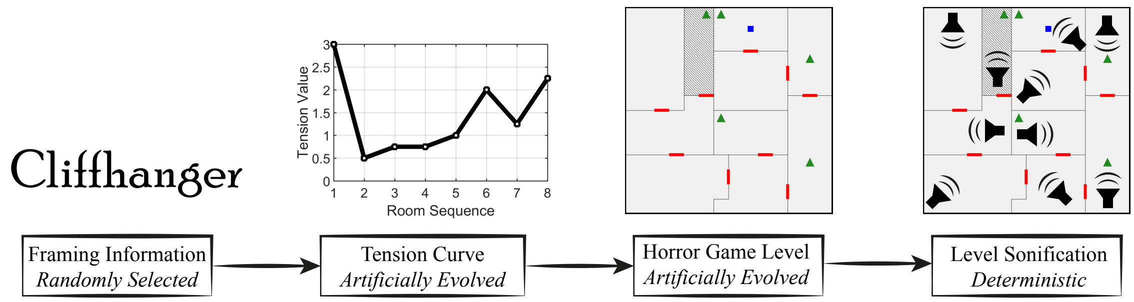


Figure 1: The creative process of the Sonancia system described here: a randomly selected tension frame is used to evaluate an evolving tension progression (the intended tension curve). Once complete, the final intended tension curve guides the evolution of a level generator which attempts to place monsters and items to match the tension curve. Finally, the evolved level and its derived tension curve are used to deterministically allocate sounds based on a crowdsourced model of aural tension. The core innovations of this paper are the framing information and the evolved tension curve (the first two modules); the level generator was first described in (Lopes, Liapis, and Yannakakis 2015a), while sonification has been improved via crowdsourced models.

Background

Sonancia attempts to blend different game facets (in this paper level design, audio and narrative): this section covers related work on blending, focusing on the audio facet.

Blending Game Facets

Digital games are a medium combining different creative facets: visuals, audio, narrative, ludus, level architecture and game-play; these facets complement each other to create specific kinds of interactive experiences (Liapis, Yannakakis, and Togelius 2014). While designing content for each facet is a creative task, blending the different facets is of utmost challenge and promise within computational creativity (Lopes and Yannakakis 2014). Game generation systems like *Angelina* (Cook, Colton, and Pease 2012) and *Game-o-matic* (Treanor et al. 2012) extensively explore how different facets of games can be combined to create interesting and thought-provoking experiences. Commercial games (designed and fine-tuned by humans) tend to blend either their rules (ludus) or level design (architecture), in the case of e.g. action-RPGs. However, suggestions for automating such blends creatively have been put forth (Gow and Corneli 2015). Blends between audio and gameplay have been explored in *AudioInSpace*, where the shooting mechanics change according to the background music, which can be hand-authored (loaded from a music library) or artificially evolved (Hoover et al. 2015). Similar studies have focused on blending audio and narrative in order to foreshadow upcoming story events via sound (Scirea et al. 2014).

The current paper builds upon and extends earlier work on *Sonancia* (Lopes, Liapis, and Yannakakis 2015a; 2015b), by allowing it to autonomously decide on an emotional progression through framing inspired by narrative structures, and by applying a crowdsourcing methodology for the emotional evaluation of sounds in the sonification audio library.

Sound and User Experience

When effectively used, audio has the potential of enhancing the player experience by fully immersing the player within a virtual world (Collins 2013). This property is especially important within the genre of horror in which particular audio patterns such as musical foreshadowing, the absence of noise, or even a rise of tempo, volume and pitch can elicit stressful experiences for players (Garner, Grimshaw, and Nabi 2010; Ekman and Lankoski 2009). These audio patterns are successful in eliciting intense affective responses if they are well interwoven with the design of the game levels. Earlier work of the authors explored how this could be achieved by sonifying levels based on a common progression of tension (Lopes, Liapis, and Yannakakis 2015a). In those studies each sound assset was given an empirical measure of how tense that particular sound was perceived, allowing the *Sonancia* system to effectively place sounds that accommodate the rise and fall of tension during play (Lopes, Liapis, and Yannakakis 2015b).

Inspired by earlier success of crowdsourcing for annotating highly subjective notions such as game aesthetics (Shaker, Yannakakis, and Togelius 2013), the previous *Sonancia* system (Lopes, Liapis, and Yannakakis 2015b) is extended via crowdsourcing of annotations on tension for sound samples. Such annotations can be used to derive more accurate data-driven computational models of tension in horror games, and offer *Sonancia* a human-verified, objective and more reliable way to select and place sounds to create spooky, tense soundscapes.

Methodology

Sonancia consists of several generative modules working as a pipeline (see Fig. 1): each generator restricts and guides the type of content which can be created in the next generative step, and with each step the content becomes more refined. The final result is a complete horror game, where players must reach a specific room within a haunted mansion while avoiding terrifying monsters along the way (see Fig. 2a). Players do not have weapons and must avoid direct confrontation with monsters; monsters thus act as an instigator of tension and fear, regardless of the player's skill.

Levels in *Sonancia* are generated via evolutionary computation guided by intended tension frames, evolved previously. Every *Sonancia* level consists of rooms connected by doors; rooms can have monsters to be avoided and the objective which must be reached to complete the level (see Fig. 2a). The level is characterized by its *critical path*, which is the shortest sequence of rooms (i.e. shortest path) between the player's starting room and the room with the objective.

The version of *Sonancia* presented in this paper consists of three different generative modules (see Fig. 1): the framing of tension to a randomly chosen narrative property, the game level generation, and the level sonification module. The details of each module are presented below.

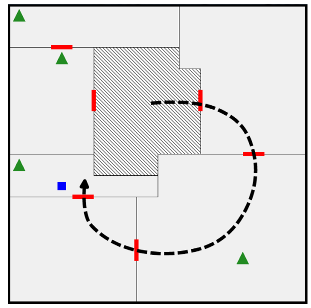
Framing Tension

To clearly explain how levels are created in *Sonancia*, it is important to firstly define how the designer intention is represented within the system. The frame for the task of horror game generation is provided by an *intended tension curve* which consists of a 2D representation of how tension rises and falls as the player progresses along the critical path (see Fig. 2b). In other words, the intended tension curve portrays the ideal player experience when going through the level.

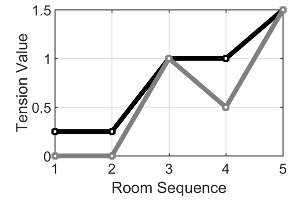
This paper specifically explores how an autonomous creative system can provide a frame to the level generation process by creating different intended tension curves. For horror games, we focus on a frame of *tension* as an amalgam of the predominant emotions within the horror genre (Ekman and Lankoski 2009): fear, anxiety and stress.

Evolving Intended Tension Curves: The intended tension curves are created via a genetic algorithm (GA), driven by one or more aesthetics of narrative progression. The GA allows for flexibility and creativity when defining the curve, but push it towards specific shapes. The tension curve is represented as as an array of values between 0 and 3 (in increments of 0.25), where the array index is the room in the order of the critical path, while each value of the array is the specific tension value. Evolution applies a roulette wheel selection mechanism with one-point crossover (Mitchell 1998). After recombination each offspring has a 20% chance of mutating, i.e. incrementing or decrementing a single value in the array by 0.25 (provided the result is within 0 and 3). The GA runs for 100 generations with a population of 100 individuals, each initialized with random tension values.

Evaluating Intended Tension Curves: Eight different fitness functions are encoded into the system, inspired by narrative structures and normalized to $[0, 1]$. The **Escalating** and **Decreasing** tension fitness rewards individuals with rooms that have a higher or lower tension value from the previous room, respectively. The **Resting Point** fitness rewards individuals with the deepest tension 'valley', while the **Surprising Moment** fitness rewards the height of the highest 'peak'. The **Cliffhanger** fitness rewards tension curves with at least one peak, where the last room's tension is higher than any of the peaks. The **Denouement** fitness gives high values to individuals if the highest peak is close to the final room (but is not the final room). **Unresolved Tension** fitness rewards consecutive rooms with the same tension. Finally the **Rising & Falling Tension** fitness is proportionate to the



(a) Example 2D level



(b) Tension curves of Fig. 2a.

(c) Playable *Sonancia* level in 3D (curated).

Figure 2: Example of a *Sonancia* "haunted manor" level in 2D (Fig. 2a) and 3D (Fig. 2c). In Fig. 2a, the room with the diagonal lines is the starting room, red rectangles are doors, green triangles are monsters, the blue square is the objective and the black arrow is the critical path (the shortest path between the starting room and objective). The critical path creates a level tension curve (grey) in Fig. 2b which must closely match the intended tension curve (black).

number of peaks in the tension curve. Among these eight fitnesses, one is chosen randomly to generate the appropriate tension frame. To increase the expressivity of generated frames, the system can also choose two fitnesses and apply an **"Or"** or **"And"** operator which sums or multiplies, respectively, the individual fitness scores.

The Level Tension Curve: Each level derives a tension curve from the distribution of monsters on the level's critical path: this process generates the *level tension curve*. Going through each room on the critical path, the level tension curve increases tension by 1 if the room contains a monster; if the room has no monster the tension decreases by 0.5 (to a minimum of 0) to simulate the players relaxing after a stressful event. Figure 2b shows the level tension curve for the level of Fig. 2a.

Level Generation

To create levels that adhere to the frame of intended tension, a search-based PCG approach was chosen (Togelius et al. 2011). The level generation process has been described in (Lopes, Liapis, and Yannakakis 2015a), but a high level description is included in this paper for the sake of completeness. The level layout is represented as an array of integers (each integral value corresponding to a room's identifier or ID), while the doors are represented by their connecting rooms' IDs and monsters or objective by the ID of the room they are in and their type. Mutations allow a gene to change the level layout (pushing walls or splitting rooms), or to add, remove and move doors, monsters or the objective. Crossover is omitted due to its disruptive nature.

Levels construct their own version of the tension curve (i.e. the *level tension curve*), and the fitness function rewards rooms which more closely match the intended tension curve. The fitness function is the average distance between level and intended tension curve (see Fig. 2b). If the level has fewer rooms on the critical path then the intended tension curve is scaled, while if it has more rooms then it receives a minimal fitness as the intended tension curve acts as a constraint on room number. In addition, the fitness function also calculates the number of unique rooms visited between the start (room ID 0) and all "dead-end" rooms and subtracts the number of rooms with no doors (as those can not be visited). More details on the evolutionary algorithm and objectives can be found in (Lopes, Liapis, and Yannakakis 2015a).

Level Sonification

Level sonification in the *Sonancia* system consists of allocating specific audio pieces within the level, based on the level tension curve. The goal of sonification is to have sounds which match the tension of the room, i.e. rooms with monsters will have scarier associated music; this is different from (Lopes, Liapis, and Yannakakis 2015b) which used sonification for suspense as the reverse of tension. *Sonancia* includes a soundbank of human-authored recordings with an average length of 7 seconds. To accurately map sound assets to specific values of tension, a crowdsourcing experiment was conducted to obtain an approximation of how tense the different sounds are compared to each other.

The Sound Library: The *Sonancia* sound library currently contains 97 different sound assets, recorded by human authors via the *FM8* (Native Instruments 2006) tool and the *Reaper* (Cuckos 2005) digital audio workstation. To maintain a large, yet feasible number of samples for crowdsourced annotation, we undersampled sounds from the library based on their "pitch" and "loudness".

According to Garner, Grimshaw, and Nabi (2010) loud (i.e. power) and high-pitched sounds tend to trigger fearful emotions. Based on this finding, we plotted (see Fig. 3) each audio asset according to the ΔDb value (loudness) and average power of frequencies above 5k (high-pitch). For the crowdsourcing experiment presented in this paper we selected the 40 sounds (out of the 97 available) with the highest average Euclidean distance between them along pitch and loudness (see Fig. 3).

Crowdsourcing Tension: A survey was conducted to derive an approximate value of reported tension to each sound asset in the library[1]. For annotating the tension value of sounds we adopt a rank-based approach due to its evidenced

[1] sonancia.institutedigitalgames.com

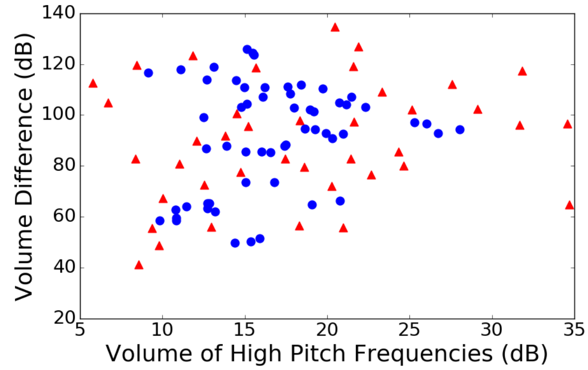


Figure 3: Scatter plot of the entire *Sonancia* sound library. High pitch frequencies are between 5 and 22 Hz, while volume difference is between the maximum and minimum dB values of the sound. Triangles and circles are, respectively, selected and unselected audio samples.

effectiveness for highly subjective notions such as affect and emotion (Yannakakis and Hallam 2011). Human annotators were presented with pairs of sounds selected randomly and were asked to report which sound in each pair is more tense via a 4-alternative forced choice questionnaire (Yannakakis and Hallam 2011). Annotators could listen to the two selected sounds as many times as they desired. At the time of writing, 452 pairs of sounds have been ranked by tension. While this is a smaller number than the 780 possible pairings, the sound pairs were randomized and thus all sounds were annotated for at least half of the possible pairings; some insight on every sound's tension properties can be gleaned even with the limited data.

The 40 sounds are ranked based on the human-annotated tension preferences. The *global order* of sound tension is derived through the pairwise preference test statistic (Yannakakis and Hallam 2011) which is calculated as $P_i = (\sum_j^N z_{ij})/N$, where z_{ij} is the tension preference score of i in the pair of sounds i and j (z_{ij} is $+1$ if sound i is preferred, -1 if sound j is preferred, and 0 if no sound is preferred or there is no annotation); N is the total number of sounds. The obtained tension preference scores P define the global order (rank) of each sound with respect to tension.

Audio Allocation and Mixing: Audio allocation consists of placing sound assets in each room of a level, based on the level tension curve and the tension preference score of each sound in the library. The system picks sounds equidistantly from the global order (in descending tension preference score) depending on the total number of rooms (not only those in the critical path). A sound is assigned to each room so that the room's tension value matches the global order of sound tension. The process starts with the most tense sound which is allocated to the room with the highest tension and it continues until no more rooms (or sounds) are available and each room has a unique sound. Higher ranked sounds with respect to tension are prioritized for rooms on the critical path. For rooms with equal tension values, the first room in the critical path gets the more tense sound.

The mixing algorithm controls how the sounds are played in the game. Audio mixing uses the player's distance from a neighbouring room to adjust the volume of the contribution from each neighbouring room's sound. This mixing rule allows players to hear sounds from neighbouring rooms, offering a sense of foreshadowing.

Experiments

This section describes results obtained from *Sonancia*'s entire process, from creating a framing of tension to generating levels based on this frame and finally sonifying the level. The goal is to evaluate, in a qualitative way, how the different generators interpret (in a tension graph, level structure or sound sequence) a frame of increasing detail created by the previous generative step in the pipeline of Fig. 1. The discussion of results assesses how accurately, for instance, the levels match the tension curves and where the limitations of one generative domain lead to a creative transformation of the other domain's data.

The system ran independently 40 times, where the framing fitnesses were selected (and often combined) by the system without human intervention. Once a framing fitness is selected, intended tension curves evolved for 100 generations in 20 independent runs; the fittest one among these runs is selected to guide level generation. Level generation performed 20 independent runs for 100 generations using the intended tension curve found previously. For brevity, we discuss the fittest individuals (tension curves, levels and soundscapes) for four chosen generated frames; these provide the most varied and interesting results. The highlighted system's frames were provided in text as such:

1. "I want an experience with a denouement."
2. "I want an experience with a cliffhanger."
3. "I want an experience with both a surprising moment and a point of rest."
4. "I want an experience with decreasing tension or a cliffhanger."

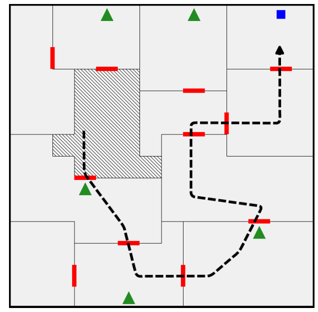
The following sections describe (in the above order) the tension curves, levels and soundscapes created following this computer-generated frames of tension.

Framing Denouement

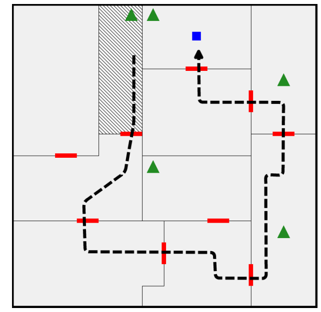
Denouement (or conclusion) is encoded aesthetically as a fitness function which rewards when the highest peak in the tension curve is near the last room (but not the last room, as that would not form a 'peak' per se). Observing the fittest intended tension curve in Figure 4c, the intended tension curve (in black) matches this specification as the highest peak (at 2.5) is on the 7th room out of 8 rooms on the critical path.

Level Generation: Figure 4a shows the fittest level for the intended tension curve discussed above; its level tension curve is shown in Fig. 4c, in grey. It is immediately obvious that the level tension curve does not match the intended one closely, although it does have a single peak at room 4 and a denouement of 4 rooms after that (rooms 5-8). The level cannot match the intended tension curve since e.g. monsters always add 1 to the tension and decay does not allow the 'constant' tension between rooms 3 and 4 or the quick drops of rooms 5 and 6. Instead, evolution attempts to balance the tradeoffs between monsters and tension decay, by adding or removing monsters in specific rooms. The result in Fig. 4a contains 3 monsters in the first three rooms after the starting one, and then no monsters until the objective room — allowing the player to relax. The biases and constraints of the level generation forced evolutionary search to interpret the intended tension curve to the best of its ability; the level tension curve does exhibit denouement, albeit lasting longer.

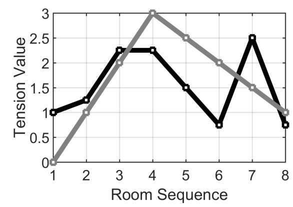
Level Sonification: Table 1a shows the distribution of different audio assets and their respective rank value within the level of Fig. 4a. It is important to note that sonification will always follow the level tension curve, to create a



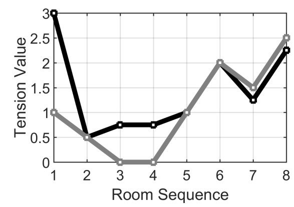
(a) Best level for Denouement



(b) Best level for Cliffhanger



(c) Tension for Denouement



(d) Tension for Cliffhanger

Figure 4: Haunted mansions and their intended and actual (level) tension curves for single aesthetics.

Room	1	2	3	4	5	6	7	8
Rank	22	16	7	1	4	10	13	19
P	0.04	0.13	0.44	0.79	0.67	0.24	0.19	0.05

(a) Denouement

Room	1	2	3	4	5	6	7	8
Rank	10	16	19	22	13	4	7	1
P	0.24	0.13	0.05	0.04	0.19	0.67	0.44	0.79

(b) Cliffhanger

Table 1: Sound selection for the Denouement (Table. 1a) and Cliffhanger (Table. 4b) levels. The sounds' corresponding rank position with respect to tension (Rank) and tension preference score (P) are also presented.

soundscape coherent to the current level. In this instance the algorithm places the highest ranked sound at the tension peak (i.e. room 4). Room 3 has a higher ranked sound than room 6 even though they have the same tension values, as it occurs first on the critical path. A game-play video of this level is available online[2].

Framing the Cliffhanger

The cliffhanger is encoded aesthetically as a fitness function which rewards tension curves with at least one peak, where the last room's tension is higher than any of the peaks (acting, thus, as the cliffhanger). The fittest intended tension curve in Figure 4d (in black) matches this specification as the tension peaks in the 6th room (with a value of 2.0) but the final room is even more tense (2.25). It should be noted that the curve starts at the highest value (3.0) in room 1; this is due to the fact that the first room does not register as a peak (peaks compare tension values with both neighbours).

Level Generation: Figure 4b shows the fittest level for the intended tension curve discussed above; its level tension curve is shown in Fig. 4d, in grey. Unlike denouement, the level tension curve closely matches the intended one for the cliffhanger aesthetic. Both curves start at the maximum possible tension (for levels, this is 1 if there is a monster in the first room) and then drop the tension in the next rooms only to increase it around rooms 5 and 6, culminating at the highest value (ignoring the first room in the intended curve) on room 8. Interestingly the level curve drops to 0 in rooms 3 and 4 as it can not maintain the near-identical tension of the intended curve (due to tension decay).

The result in Fig. 4b has 4 monsters on the critical path, distributed near the start and end of this path. This causes an initial tense moment for the players when they start the level, then lets them relax with 3 empty rooms, reach a climax after two monsters and release some tension with the next-to-last-room only to find a monster in the room with the objective.

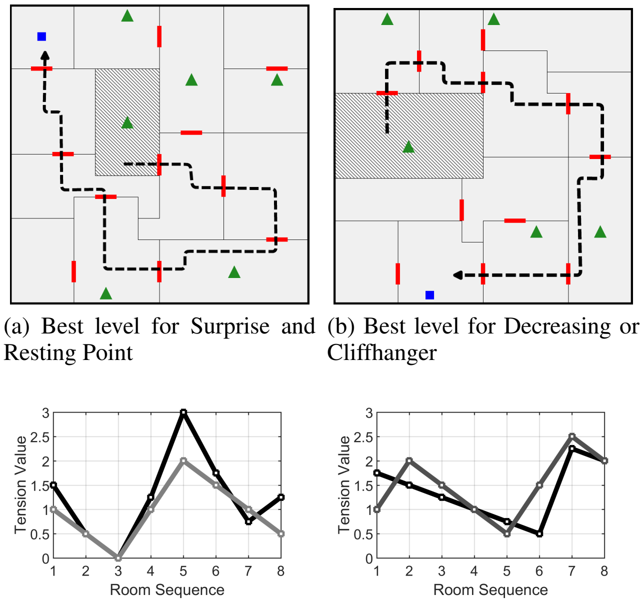
Level Sonification: Table 1b contains the sounds allocated along the critical path of the level in Fig. 4b). As the level tension curve closely matches the intended curve, sonification largely matches the original frame as well. While rooms 2 to 5 have sounds with a low global rank value, this changes swiftly with tense sounds which culminate to the most tense sound in the last room. A game-play video of this level is available online[3].

Frame of Surprising Moments and Resting Points

When combining fitness functions, the "and" combination forces both fitnesses to have high scores: in this case, the surprising moment aesthetic rewards high 'peaks' while the resting point aesthetic rewards deep 'valleys'. Indeed, both aesthetics are present in the intended tension curve of Fig. 5c as it exhibits the highest peak (height of 3) and the lowest possible valley (depth of 3, considering the tallest adjacent peak). The aggressive changes in tension were expected, as both fitnesses directly reward high peaks and deep valleys; their combination unsurprisingly causes tension to soar from a value of 0 to 3 within the span of two rooms. In many other runs, the fittest individuals contained adjacent rooms with tension values of 0 and 3 (or vice versa).

[2]https://youtu.be/IJQFqxfHqY8

[3]https://youtu.be/z5R12NPVVFA



(a) Best level for Surprise and Resting Point (b) Best level for Decreasing or Cliffhanger

(c) Tension for Surprise and Resting Point (d) Tension for Decreasing or Cliffhanger

Figure 5: Haunted mansions and their intended and actual (level) tension curves for combined aesthetics.

Room	1	2	3	4	5	6	7	8
Rank	7	16	22	10	1	4	13	19
P	0.44	0.13	0.04	0.24	0.79	0.67	0.19	0.05

(a) Surprising Moments and Resting Points

Room	1	2	3	4	5	6	7	8
Rank	16	4	10	19	22	13	1	7
P	0.13	0.67	0.24	0.05	0.04	0.19	0.79	0.44

(b) Decreasing Tension or a Cliffhanger

Table 2: Sound selection for the Surprising Moments and Resting Points (Table 2a) and Decreasing Tension or a Cliffhanger (Table 2b) levels. The sounds' corresponding rank position with respect to tension (Rank) and tension preference score (P) are also presented.

Level Generation: Figure 5a shows the fittest level for the intended tension curve discussed above; its level tension curve is shown in Fig. 5c, in grey. The level tension curve matches the intended tension curve as closely as possible given the constraints of the way it is computed. The fact that each room can have only one monster (which increases tension by 1) causes the peak of room 5 after the resting point in room 3 to have lower tension values than the intended. The structure of the level tension curve retains both a resting point at room 3 and the surprising moment at room 5, and thus matches the provided frame. Of interest is the observation that unlike the intended curve, the last room in the level does not contribute to an increase in tension since adding a monster there (getting the tension value to 2) would cause more deviation from the intended value (1.25).

The level of Fig. 5a has 3 monsters on the critical path, placed primarily midway to the objective. This yields a high spike (surprising moment) in room 5 after encountering two monsters. The starting room also has a monster, in order to let players relax in the next two rooms and thus reach the resting point (before more stressful events) in room 3.

Level Sonification: Table 2a contains the sounds allocated along the critical path of the level in Fig. 5a. Obviously, the most tense sound is placed on the surprising moment (room 5) which matches both the intended and the level tension curve; similarly, the resting point has the least tense sound as per the provided frame. Due to the similarity of the level curve with that of Fig. 4c, tense sounds are allocated in a somewhat similar fashion with a slight change in the first rooms. A game-play video of this level is available online[4].

Framing Decreasing Tension or a Cliffhanger

Combining fitness functions with an "or" in this system adds the two fitness scores together. This will still reward the presence of both features but since it is less aggressive than multiplying the scores (as in "and"), it may reward either fitness equally. The fittest tension curve in Fig. 5d, for instance, does not have the cliffhanger pattern (although partially it does exhibit a peak in room 7) but has predominantly a decreasing tension. The cliffhanger and decreasing tension are conflicting objectives, as the former rewards an increase in tension both for the presence of a peak and for the final room. Therefore, the intended curve in Fig. 5d attempts to balance between the two by predominantly having a decreasing tension while also having a peak (which is rewarded, partially, by the cliffhanger fitness). Thus, the less aggressive search of the "or" operator is demonstrated.

Level Generation: Figure 5b shows the fittest level for the intended tension curve discussed above; its level tension curve is shown in Fig. 5d, in grey. The level tension curve matches the intended tension curve except that the gradient of the tension decay is different: this causes evolution to use two rooms (6 and 7) to increase the tension in order to match the tension value of room 7 (2.25 in the intended curve and 2.5 in the actual one). Interestingly, despite the expected differences when the level generator interprets the intended frame (e.g. an increase in tension at room 2), the aesthetics match between intended and level tension curve. The level tension curve predominantly has decreasing tension, with no cliffhanger but at least one peak (thus fulfilling one of the requirements for a cliffhanger).

The level of Fig. 5b has 4 monsters on the critical path, placed at the start and towards the end of the critical path. The two monsters in the first and second rooms trigger a very tense experience to the player, but the decreasing tension aesthetic allows them to relax for the next 4 rooms before facing two more monsters in rooms 6 and 7. The room with the objective does not have a monster, affording some relaxation to the player.

Level Sonification: Table 2b shows how sounds were allocated within the level and their respective values. Compared to the other sonification results, this soundscape spreads highly tense sounds throughout the level rather than concentrating them in a specific section. Interestingly, the level tension curve is unique compared to the other cases as no room has a tension value of 0. For instance, room 2 has the second highest ranked sound, but is surrounded by less tense sounds, while the most tense sound is reserved for the climax (i.e. room 7). A game-play video of this level is available online[5].

Discussion

The results highlighted four example tension frames which were associated with one or multiple fitness dimensions. The results showed that the intended tension curves created by the system matched the patterns in the narrative structures they were based on. An exception was when conflicting fitnesses were combined with the "or" operator, where one fitness could dominate the other (earning the operator its name). The generated levels in many cases matched the intended curve (if not value-for-value) but the limitations of the level tension curve calculation could cause deviations (e.g. in the case of denouement). At a high-level, all generated levels exhibited the intended aesthetics of each frame.

Observing results with other fitness dimensions of framing, we found that *Escalating*, *Decreasing* and *Unresolved Tension* fitnesses created the least variability in the tension curves. This was expected, as these fitnesses reward small incremental changes in the tension or no changes (for Unresolved Tension). Both the *Surprising Moment* and *Resting Point* fitnesses created more variations in the tension curves but both showed similar patterns: a drastic change of tension (from 0 to 3 or vice versa) between two adjacent rooms (similar to Fig. 5c). This pattern is impossible to replicate in the levels, leading to more free-form interpretation of the intended curve by the level generator. An interesting emergent solution to attain less aggressive tension changes was when fitnesses were combined: for instance, combining any fitness with the Escalating or the Decreasing fitness yielded curves with smoother changes in tension. Due to a less strict evaluation formula, the *Denouement*, *Cliffhanger* and *Rising & Falling Tension* created the most diverse curves. Peaks very often varied in tension, and in some cases the entire curve would have low values of tension, or only high values.

The additional modules of the *Sonancia* pipeline (highlighted in Fig. 1) contribute to the creativity of the system in two core ways: *framing information* and *interpretation*. Framing information (as desired narrative structures) allow the generator to describe in human language its intent; the fitness function associated with each narrative structure al-

[4] `https://youtu.be/P2HkGr719f0`

[5] `https://youtu.be/JnFli_F-r38`

lows the system to *appreciate* whether it has achieved this intent. The fact that the initial frame is chosen randomly is a current limitation of the system; its creativity could be strengthened if the inspiration for a narrative structure comes from elsewhere (e.g. a newspaper article). Interpretation in *Sonancia* is strengthened by extending the pipeline to include generated tension curves which guide the level generator, which in turn guides the sonification process: as the level of detail of the creative artifact increases from an abstract frame to a playable game, the generators must creatively interpret the guidelines of the previous generative step in order to satisfy them while still obeying the limitations of their own level of detail (e.g. the structural requirements of a level). This requires a degree of *imagination* from each module in transforming high-level directives into higher-detail artifacts. Finally, as the final game levels are always playable and contain the necessary components for horror gameplay, Sonancia has the necessary *skill* and thus completes the creative tripod exhibited by creative systems (Colton 2008).

Conclusions

This paper presented a system capable of creating and combining different structures of tension influenced by narrative concepts, then transforming them into horror levels and their accompanying soundscapes. Several dimensions of tension framing structure were developed and tested; the four example generated frames highlighted the process from frame conceptualization to level generation and finally to sonification. Demonstrations of the sonification of all the example levels in this paper can be found online[6].

Acknowledgments

The authors would like to thank all participants of the crowdsourcing experiment. This work was supported, in part, by the FP7 Marie Curie CIG project AutoGameDesign (project no: 630665) and the Horizon 2020 project CrossCult (project no: 693150)

References

Collins, K. 2013. *Playing with sound: a theory of interacting with sound and music in video games*. MIT Press.

Colton, S.; Charnley, J.; and Pease, A. 2011. Computational creativity theory: The face and idea descriptive models. In *Proceedings of the Second International Conference on Computational Creativity*.

Colton, S. 2008. Creativity vs. the perception of creativity in computational systems. In *Papers from the AAAI Spring Symposium on Creative Intelligent Systems*.

Cook, M., and Colton, S. 2014. A rogue dream: Automatically generating meaningful content for games. In *Proceedings of the AIIDE workshop on Experimental AI & Games*.

Cook, M.; Colton, S.; and Pease, A. 2012. Aesthetic considerations for automated platformer design. In *Proceedings of the Artificial Intelligence and Interactive Digital Entertainment conference*.

[6]`https://goo.gl/Qshvqv`

Ekman, I., and Lankoski, P. 2009. Hair-raising entertainment: Emotions, sound, and structure in silent hill 2 and fatal frame. *Horror video games: Essays on the fusion of fear and play* 181–199.

Garner, T.; Grimshaw, M.; and Nabi, D. A. 2010. A preliminary experiment to assess the fear value of preselected sound parameters in a survival horror game. In *Proceedings of the 5th Audio Mostly Conference: A Conference on Interaction with Sound*, 10. ACM.

Gow, J., and Corneli, J. 2015. Towards generating novel games using conceptual blending. In *Eleventh Artificial Intelligence and Interactive Digital Entertainment Conference*.

Grace, K.; Gero, J.; and Saunders, R. 2012. Representational affordances and creativity in association-based systems. In *Proceedings of the International Conference on Computational Creativity*.

Hoover, A. K.; Cachia, W.; Liapis, A.; and Yannakakis, G. N. 2015. AudioInSpace: Exploring the creative fusion of generative audio, visuals and gameplay. In *Proceedings of the EvoMusArt conference*. Springer. 101–112.

Johnson, D., and Ventura, D. 2014. Musical motif discovery in non-musical media. In *Proceedings of the International Conference on Computational Creativity*.

Krzeczkowska, A.; El-Hage, J.; Colton, S.; and Clark, S. 2010. Automated collage generation – with intent. In *Proceedings of the International Conference on Computational Creativity*.

Liapis, A.; Yannakakis, G. N.; and Togelius, J. 2014. Computational game creativity. In *Proceedings of the International Conference on Computational Creativity*, 285–292.

Lopes, P., and Yannakakis, G. N. 2014. Investigating collaborative creativity via machine-mediated game blending. In *Proceedings of the Artificial Intelligence and Interactive Digital Entertainment conference*.

Lopes, P.; Liapis, A.; and Yannakakis, G. N. 2015a. Sonancia: Sonification of procedurally generated game levels. In *Proceedings of the 1st Computational Creativity and Games Workshop*.

Lopes, P.; Liapis, A.; and Yannakakis, G. N. 2015b. Targeting horror via level and soundscape generation. In *Eleventh Artificial Intelligence and Interactive Digital Entertainment Conference*.

Mitchell, M. 1998. *An introduction to genetic algorithms*. MIT press.

Scirea, M.; Cheong, Y.-G.; Nelson, M. J.; and Bae, B.-C. 2014. Evaluating musical foreshadowing of videogame narrative experiences. In *Proceedings of the Audio Mostly: A Conference on Interaction With Sound*. ACM.

Scirea, M.; Barros, G. A. B.; Shaker, N.; and Togelius, J. 2015. SMUG: Scientific music generator. In *Proceedings of the International Conference on Computational Creativity*.

Shaker, N.; Yannakakis, G.; and Togelius, J. 2013. Crowd-sourcing the aesthetics of platform games. *IEEE Transactions on Computational Intelligence and AI in Games* 5(3).

Togelius, J.; Yannakakis, G. N.; Stanley, K. O.; and Browne, C. 2011. Search-based procedural content generation: A taxonomy and survey. *Proceedings of the Computational Intelligence and Games Conference* 3(3):172–186.

Treanor, M.; Blackford, B.; Mateas, M.; and Bogost, I. 2012. Game-o-matic: Generating videogames that represent ideas. In *Proceedings of the FDG workshop on Procedural Content Generation in Games*, 11. ACM.

Yannakakis, G. N., and Hallam, J. 2011. Ranking vs. preference: a comparative study of self-reporting. In *Affective computing and intelligent interaction*. Springer. 437–446.