Affective Expression in Computer Generated Music and its Effect on Player Experience

Marco Scirea

IT University of Copenhagen, Copenhagen 2300, DK,
msci@itu.dk,
WWW home page: itu.dk/people/msci

Abstract. In games, unlike in traditional linear storytelling media such as novels or films, narrative events unfold in response to the player input. Therefore, the music composer in an interactive environment needs to create music that is dynamic and non-repetitive. We investigate how to express emotions and moods in music and how to apply this research to improve player experience in games. The main novelty in our approach, compared to most algorithmically generated music research, is our focus on affective and cognitive modelling, coupled with real-time adjustment of the music. This focus on the emotional meaning that procedurally generated music should express has also been identified by Collins as one of the lacking features that prevent procedurally generated music to be more widely used in the game industry [1].

Keywords: affective computing, generative music

1 Problem Statement

We aim to fill what we believe is the gap that's holding generative procedural music generation back: emotion expression. A number of works have been published in the area of affect, semiotics and mood-tagging ([2],[3],[4]) but our focus lies in the real-time generation of background music capable of expressing moods. Our research focuses on investigating the expression of moods and the affective effect this music can have on the listener while applying this research to games.

The final objective is to create a system where, by using a emotional model of the player, we would be able to identify the player's emotional state and be able to reinforce or manipulate it through the use of mood-expressive music to improve user experience. What this could achieve is the creation of better immersive experiences (reinforcement of current emotional state and play-style) and the development of tools/models to help designers create experiences where the players are put in a specific emotional state.

A number of challenges arise from this objective, such as how to validate our mood expression model, improve our generator to be able to include harmony and melody, make the generated music more interesting and creating a cognitive model of the player, just to name a few. First we are going to validate the theory we used until now and better tune it to increase the mood recognition rate, then we will harmony generation in our generator. We are thinking about training a Markov chain model by using a database of chord successions which could consider, apart from the current chord, one or two previous ones.

Our current generator doesn't consider harmony, so we can't express these; still we will soon integrate harmony and melody generation in our framework, hopefully opening up even more possibilities for our research. Once we have improved our music generation to ideally being able to express easily identifiable moods and creating interesting music, we will start working on the cognitive model of the player to find ways to extrapolate his/her emotional state. This will be integrated into an affective loop where music generation is used as part of an approach to player-adaptive games through content generation. We would firstly focus on one specific game genre and, time permitting, expand our model to include more genres, making it more general. We would also like to continue our work on narrative cues expression through music, even if this direction of the research is not the main focus of the project.

2 State of the art

2.1 Procedurally generated music

Procedural generation of content is a booming field that has seen a tremendous growth lately. Applications can be: creating simple sound effects, game levels, entire game worlds, and more. While a good number of games use some sort of procedural music structure, there are different approaches (or degrees), Wooller et al. distinguish two of them: transformational algorithms and generative algorithms[5]. Transformational algorithms act upon an already prepared structure, for example by having the music recorded in layers that can be added or subtracted at a specific time to change the feel of the music. (The Legend of Zelda: Ocarina of Time is one of the earliest games that use this approach).

Generative algorithms instead create the musical structure themselves, which leads to a higher degree of difficulty in keeping the music consistent with the game events, and generally require more computing power as the musical materials have to be created in real-time. An example of this approach can be found in Spore: the music written by Brian Eno was created with Pure Data in the form of many small samples that created the soundtrack in real time.

2.2 Emotions and moods

Emotions have been extensively researched within psychology, although their nature (and what constitutes the basic set of emotions) is still widely debated. Lazarus argues that *"emotion is often associated and considered reciprocally influential with mood, temperament, personality, disposition, and motivation"* [6]. Our approach is therefore to produce scores with an identifiable mood, and in so doing, induce an emotion response from the game player.

Affect is generally considered to be the experience of feeling or emotion. It is largely believed that affect is post-cognitive; emotion arises only after an amount of cognitive processing has been accomplished. With this assumption, every affective reaction (e.g., pleasure, displeasure, liking, disliking) results from *"a prior cognitive process that makes a variety of content discriminations and identifies features, examines them to find value, and weighs them according to their contributions"* [7]. Another view is that affect can be both pre- and post-cognitive, notably [8]; thoughts are created by an initial emotional response that then leads to producing affect.

Mood is an affective state. However, while an emotion generally has a specific object of focus, moods tends to be more unfocused and diffused [9]. [10] say that mood *"involves tone and intensity and a structured set of beliefs about general expectations of a future experience of pleasure or pain, or of positive or negative affect in the future"*. Another important difference between emotions and moods is that moods, being diffused and unfocused, can last much longer (as also remarked by [11]).

In this paper, we focus on moods instead of emotions, for we expect that in games – where the player listens to the background music for a longer time, that a particular emotion is induced by the mood – and moods are more likely to be remembered by the players after their game-play. In addition, they are easier for game designers to integrate, since they represent longer-duration sentiments suitable for segments of game play.

2.3 Music mood taxonomy

The set of adjectives that describe music mood and emotional response is immense and there is no accepted standard vocabulary. For example, in the work of [12], the emotional adjective set includes *Gloomy, Serious, Pathetic* and *Urbane*.

[13] proposed a model of affect based on two bipolar dimensions: *pleasant-unpleasant* and *arousal-sleepy*, theorising that each affect word can be mapped into this bi-dimensional space by a combination of these two components. [14] applied Russell's model to music using as the dimensions of *stress* and *valence*; although the names of the dimensions are different from Russell's their meaning is the same. Also, we find different terms among different authors (e.g. [15, 16]) for apparently the same moods. We will use the terms *valence* and *arousal*, as they are the most commonly used in affective computing research.

Affect in music can in this way be divided into the four clusters based on the dimensions of valence and arousal: *Anxious/Frantic* (Low Valence, High Arousal), *Depression* (Low Valence, Low Arousal), *Contentment* (High Valence, Low Arousal) and *Exuberance* (High Valence, High Arousal). These four clusters have the advantage of being explicit and discriminable; also they are the basic music-induced moods as described in [17, 18].

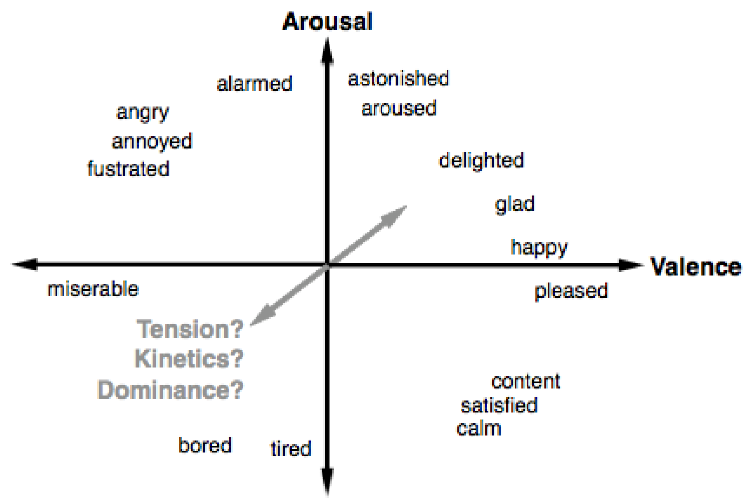

Fig. 1. The Valence-Arousal space, labelled by Russell's direct circular projection of adjectives [13]. The figure includes the projected third affect dimension: "tension", "kinetics", "dominance". In our study we have not considered this third dimension since it is still not well defined.

3 Current results and Methodology

We have created a first prototype of a Moody Music Generator [19], which can express different moods in an unstructured, semi-randomic ambient music. We have conducted multiple studies [20][21][22] to characterize it's control parameters and how effectively the moods expressed by the music can be recognized by the listener. In the first study we had some interesting results regarding emotional adjectives: there doesn't seem to be a consensus on the semantic meaning of these words and, moreover, correlations between different affective words seem to emerge. The study gave us some first encouraging results but also made us aware of the problems in our methodology.

In response we designed a new open-ended study that, rather than directly attempting to validate that our two control parameters represent arousal and valence, crowd-sourced labels characterising different parts of this two-dimensional control space. Our aim was to characterise perception of the generators expressive space, without constraining listeners responses to labels specifically aimed at validating the original arousal/valence motivation. Subjects were asked to listen to clips of generated music over the Internet, and to describe the moods with free-text labels. We find that the arousal parameter does roughly map to perceived arousal, but that the nominal valence parameter has strong interaction with the arousal parameter, and produces different effects in different parts of the control space. We believe that this characterisation methodology is general and could be used to map the expressive range of other parameterisable generators. This study has returned some positive results, yet suggests that we need to refine our mood expression model, especially on expressing the valence axis.

Currently we have implemented a new AI system for music generation, which creates more interesting and musically complex music that might influence our current affective state expression theory. The objectives of our generator are: (i) to express different affective states using a variety of AI techniques; (ii) to generate such music in real-time and (iii) to react in real-time to external stimuli. The architecture is comprised of three main components: the composition generator, the real-time affective music composer and an archive of compositions. A novel feature of our approach is the separation of composition and affective interpretation: the system creates abstractions of music pieces (what we call compositions) and interprets these compositions in real-time to achieve the desired affective expression while also introducing stochastic variations. In terms of composition generation, we present a novel combination of Evolutionary Computation techniques to evolve melodies: the Feasible/Infeasible two population method (FI-2POP [23]) and Multi Objective Optimization [24].

4 Future work

We are currently conducting an evaluation study on the music generation technique developed for our new generator. Soon we'll also study the affect expression capabilities of our generator, as our theory will probably need to be adjusted to the higher complexity of the music produced.

Soon after we plan to decide a game to integrate with the generator, and start building an affective model of the player based on that game. The affect model should be able to tell us the emotional state of the player from in-game metrics and also give us information on how we can influence the player's state through affective expressive music.

5 Contributions

This study will be (and already is) a contribution to the field of music generation: while the field is very active in the generation of music to emulate a specific style, the creation of accompaniments to a melody, there is very little research that investigates the expression (and manipulation) of affective states through procedurally generated music. We will also contribute to the field of procedural content generation, in fact our research is strongly connected with the concept expressed by Yannakakis and Togelius of Experience-Driven Procedural Content Generation (EDPC) [25], where the content itself is tailored to the player in an attempt to create highly personalized content to improve player experience. We will create the first player-adaptive affective game music generator. We believe that this innovative research and our unique approach to it might also be interesting for other fields, such as musicology, human-computer interaction and computer science in general.

References

1. Collins, K.: An introduction to procedural music in video games. Contemporary Music Review **28**(1) (2009) 5–15
2. Birchfield, D.: Generative model for the creation of musical emotion, meaning, and form. In: Proceedings of the 2003 ACM SIGMM Workshop on Experiential Telepresence. (2003) 99–104
3. Eladhari, M., Nieuwdorp, R., Fridenfalk, M.: The soundtrack of your mind: mind music-adaptive audio for game characters. In: Proceedings of Advances in Computer Entertainment Technology. (2006)
4. Livingstone, S.R., Brown, A.R.: Dynamic response: Real-time adaptation for music emotion. In: Proceedings of the 2nd Australasian Conference on Interactive Entertainment. (2005) 105–111
5. Wooller, R., Brown, A.R., Miranda, E., Diederich, J., Berry, R.: A framework for comparison of process in algorithmic music systems. In: Generative Arts Practice 2005 — A Creativity & Cognition Symposium. (2005)
6. Lazarus, R.S.: Emotion and Adaptation. Oxford University Press (1991)
7. Brewin, C.R.: Cognitive change processes in psychotherapy. Psychological review **96**(3) (1989) 379
8. Lerner, J.S., Keltner, D.: Beyond valence: Toward a model of emotion-specific influences on judgement and choice. Cognition & Emotion **14**(4) (2000) 473–493
9. Martin, B.A.: The influence of gender on mood effects in advertising. Psychology & Marketing **20**(3) (2003) 249–273
10. Batson, C.D., Shaw, L.L., Oleson, K.C.: Differentiating affect, mood, and emotion: Toward functionally based conceptual distinctions. (1992)
11. Beedie, C., Terry, P., Lane, A.: Distinctions between emotion and mood. Cognition & Emotion **19**(6) (2005) 847–878
12. Katayose, H., Imai, M., Inokuchi, S.: Sentiment extraction in music. In: Proceedings of the 9th International Conference on Pattern Recognition. (1988) 1083–1087
13. Russell, J.A.: A circumplex model of affect. Journal of Personality and Social Psychology **39**(6) (1980) 1161–1178
14. Thayer, R.E.: The Biopsychology of Mood and Arousal. Oxford University Press (1989)
15. Wundt, W.: Outlines of psychology. Springer (1980)
16. Schlosberg, H.: Three dimensions of emotion. Psychological review **61**(2) (1954) 81
17. Kreutz, G., Ott, U., Teichmann, D., Osawa, P., Vaitl, D.: Using music to induce emotions: Influences of musical preference and absorption. Psychology of Music **36**(1) (2008) 101–126
18. Lindström, E., Juslin, P.N., Bresin, R., Williamon, A.: "Expressivity comes from within your soul": A questionnaire study of music students' perspectives on expressivity. Research Studies in Music Education **20**(1) (2003) 23–47
19. Scirea, M.: Mood dependent music generator. In: Proceedings of Advances in Computer Entertainment. (2013) 626–629
20. Scirea, M., Cheong, Y.G., Bae, B.C.: Mood expression in real-time computer generated music using pure data. In: Proceedings of the International Conference on Music Perception and Cognition. (2014)
21. Scirea, M., Cheong, Y.G., Bae, B.C., Nelson, M.: Evaluating musical foreshadowing of videogame narrative experiences. In: Proceedings of Audio Mostly 2014. (2014)

22. Scirea, M., Nelson, M.J., Togelius, J.: Moody music generator: Characterising control parameters using crowdsourcing. In: Evolutionary and Biologically Inspired Music, Sound, Art and Design. Springer (2015) 200–211
23. Kimbrough, S.O., Koehler, G.J., Lu, M., Wood, D.H.: On a feasible–infeasible two-population (fi-2pop) genetic algorithm for constrained optimization: Distance tracing and no free lunch. European Journal of Operational Research **190**(2) (2008) 310–327
24. Deb, K.: Multi-objective optimization using evolutionary algorithms. Volume 16. John Wiley & Sons (2001)
25. Yannakakis, G.N., Togelius, J.: Experience-driven procedural content generation. IEEE Transactions on Affective Computing **2**(3) (2011) 147–161