

Affect-Conditioned Ambient Music Synthesis

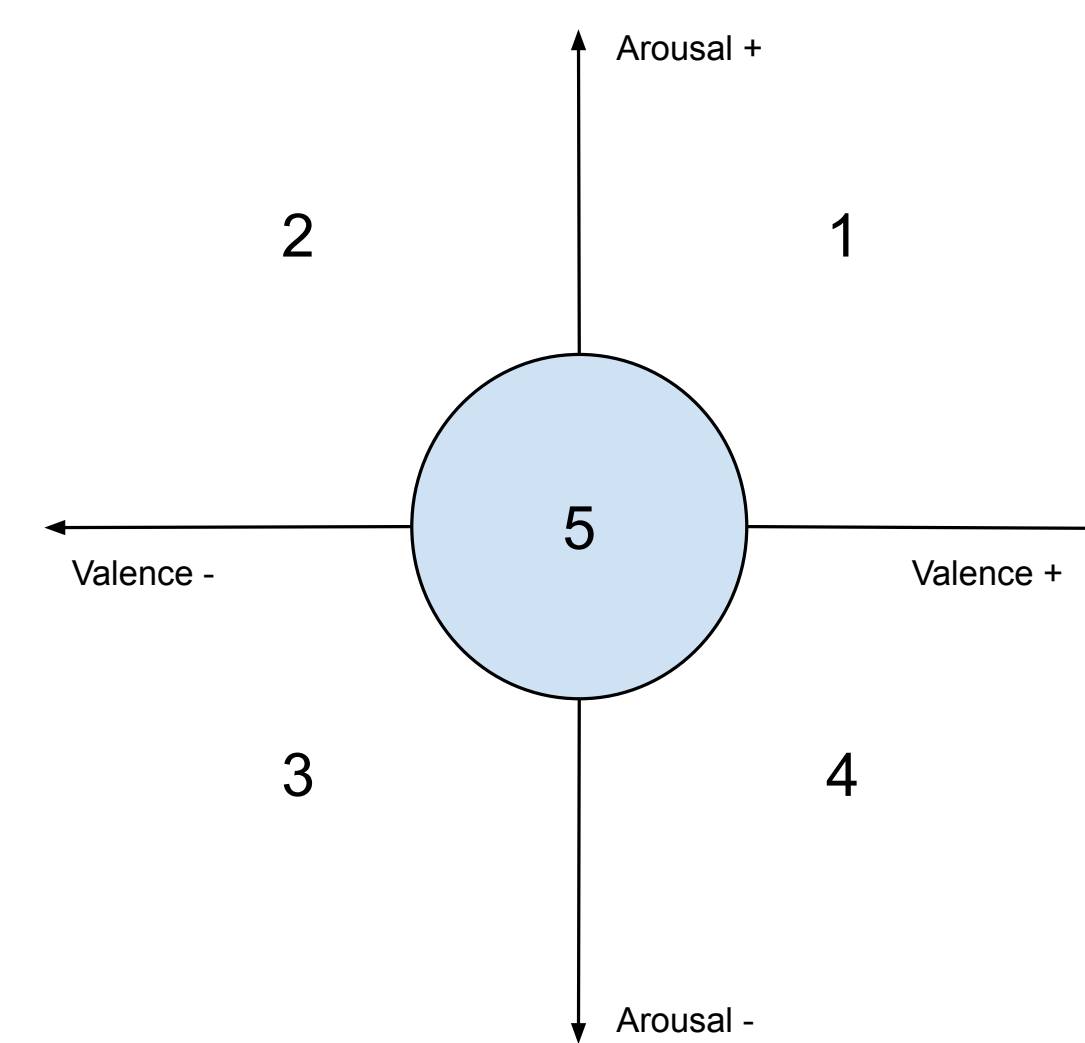
VEDANT KALBAG
Music Informatics Group
vedant.kalbag@gatech.edu



PROBLEM STATEMENT AND MOTIVATION

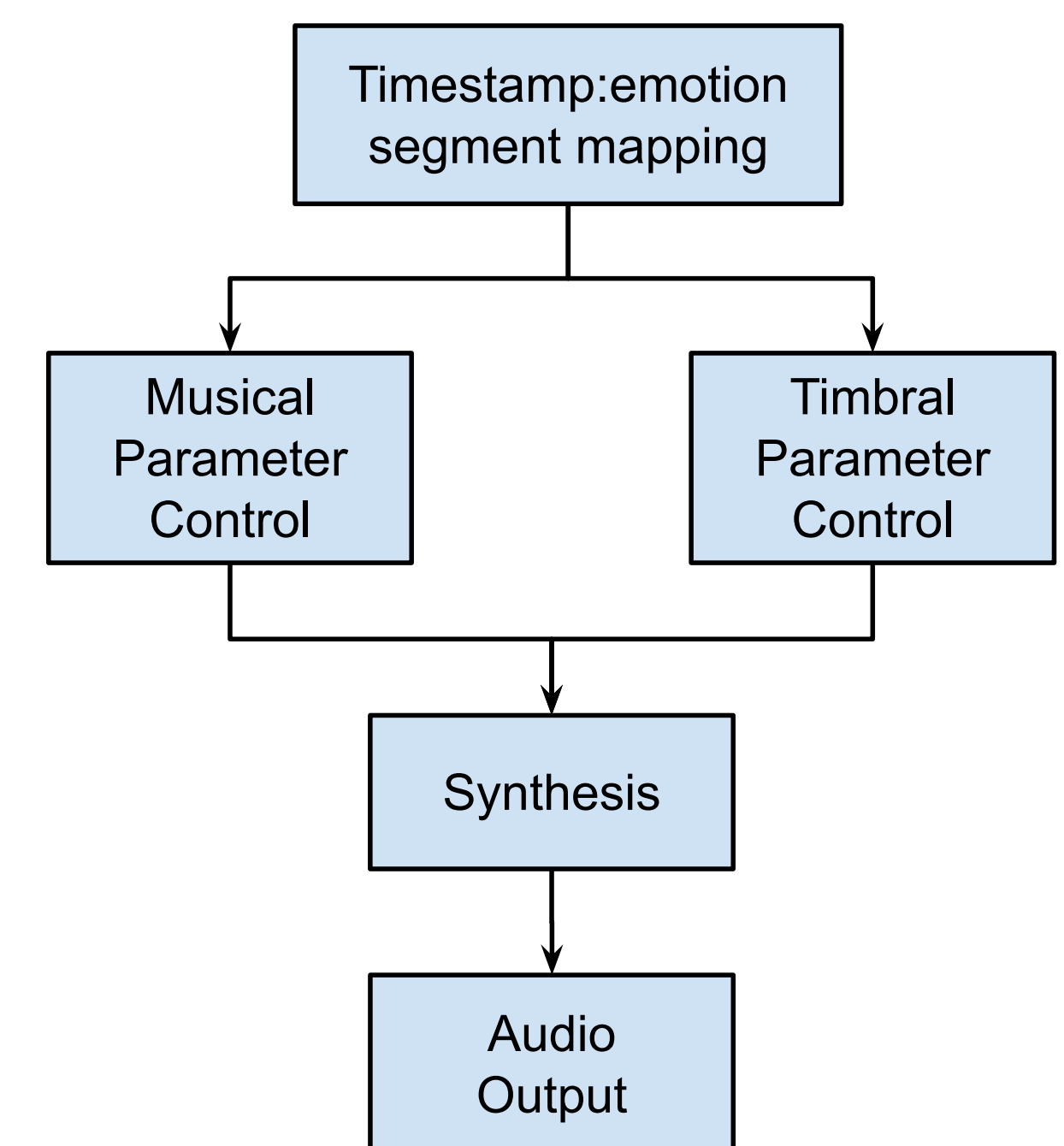
- **GOAL:** To create an open source tool for synthesis of ambient music using emotional cues
- **APPLICATIONS:** Background music generation for videos, audiobooks, eBooks, video games
- **INPUT:** emotion segment to timestamp mapping (eg. {time1:segment1, time2:segment2}, etc.)
- **OUTPUT:** audio file of desired duration with smooth transitions between the emotional segments
- Emotional music synthesis is the computerized generation of music which has a recognizable emotion, for the purpose of setting a mood
- Background music can set the mood or level of intensity of a scene for storytelling applications, whether for audiobooks or for movies

The emotion content will be controlled through manipulation of both musical parameters and timbral parameters during synthesis, based on the segment selected from the figure below



Circumplex plane mapping for emotional segments

METHOD



Overall system flowchart

MUSICAL PARAMETER CONTROL

Musical parameters used in related work:

- **Rhythmic roughness** (rate of change of note durations)
- **Number of simultaneous notes**
- **Melodic/pitch range**
- **Loudness and loudness variation**
- **Melodic direction**
- **Tonality**

The parameters in **BOLD** have been selected for use in the system for ambient synthesis, based on which parameters have the largest impact on **perceived emotion content**

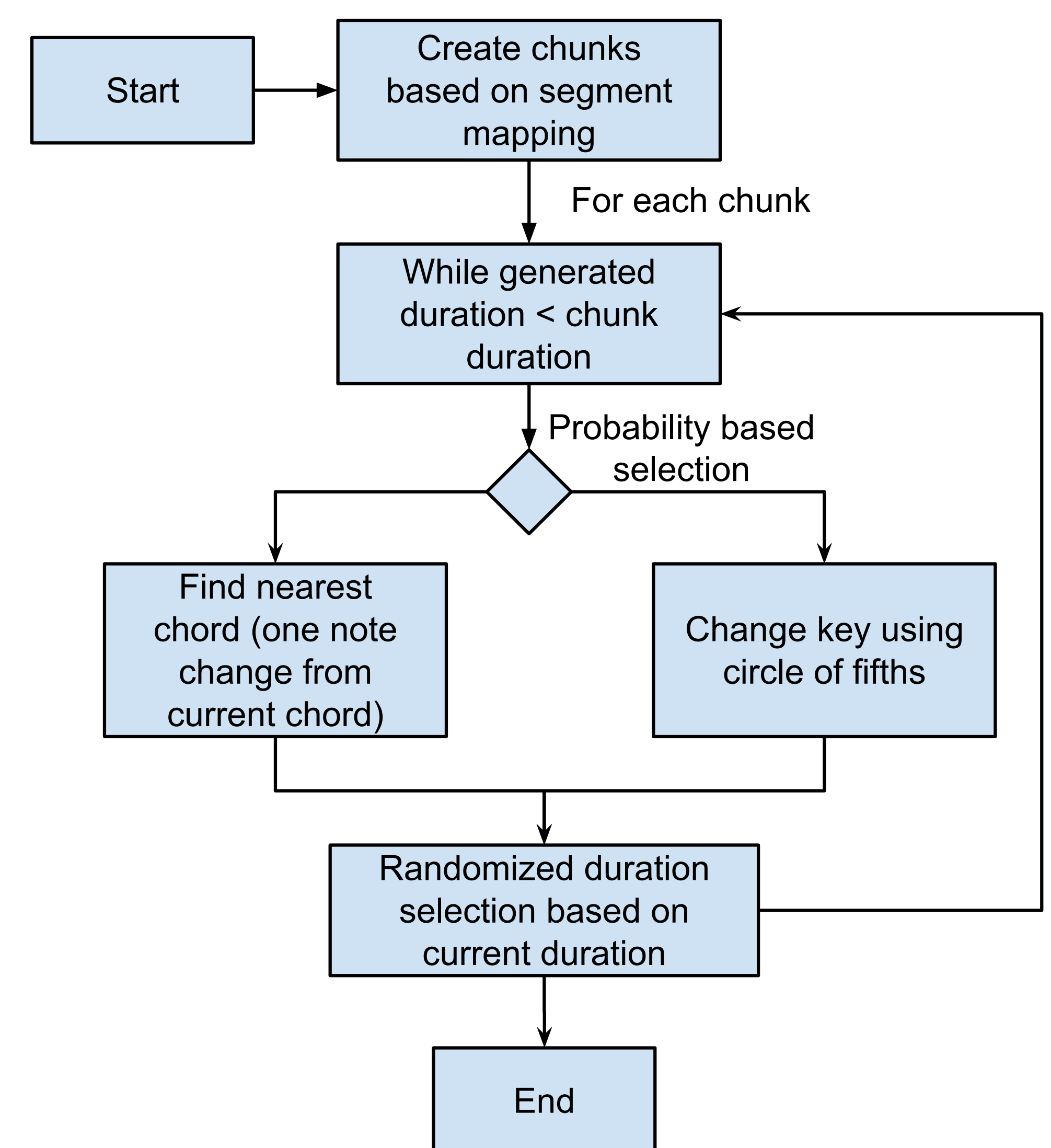
PROCESSING ALGORITHM:

1. The mapping provided by the user is broken into chunks of (time1:time2),(emotion1,emotion2)
2. For each chunk, a list of chords and durations are generated
3. These chunks are recombined into a list of chords and durations that represent the entire piece

Important considerations:

- The durations are randomly selected based on a transition matrix of "lengths" (VS, S,M,L,VL), and a random selection of actual duration based on the selected "length"
- The number of simultaneous notes are selected using a transition matrix, picking from a range of 1-5
- The chords are generated based on number of notes to be used (eg. (1), (1,5), (1,3,5), (1,3,5,9), (1,3,5,9,13/-3))
- Octaves are also chosen based on current octave and a transition matrix

These chords and durations are used with the timbral parameters for synthesis



Musical parameter control block diagram

TIMBRAL PARAMETER CONTROL

Timbral parameters used in related work:

- **ADSR**
- **Waveform types**
- **Number of voices, detune amount**
- **Filters**
- **Effects**

The parameters in **BOLD** have been selected for use in the system for ambient synthesis, based on which parameters have the largest impact on **perceived emotion content**

The presets for each emotion are manually created based on perceived emotional content of the sound

The system has multiple presets for each emotional segment, and these are chosen at random

Using the chunks created in the previous step, the beginning and ending emotional segment are compared, and if they are different, the parameter values are interpolated between a preset chosen for the initial and final emotional segments

Each chord that is to be played is given a set of timbral parameters that are sampled from this interpolation



Example synth with similar parameters for control (used for tone matching)

FUTURE WORK

- **FINE TUNING:** The system needs to be fine-tuned by adjusting the probabilities within the various transition matrices
- **EVALUATION:** To be done using ValNet and ArNet to predict valence and arousal values to verify whether the desired emotional segment is being generated
- **APPLICATION OF SYSTEM:** The system will be built into a sentiment analysis pipeline to test its application with an eBook
- **LISTERING STUDY:** A listening study will be conducted to evaluate the efficacy of the system