With Bolts of Melody!
Songwriting with ALYSIA & Emily Dickinson

Christopher Cassion
Georgia Institute Of Technology
Atlanta, GA
ccassion3@gatech.edu

Margareta Ackerman
Santa Clara University
Santa Clara, CA
ackermanmaya@gmail.com

David Loker
Ayzenberg Group
Pasadena, CA
dloker@gmail.com

Joshua Palkki
CSU, Long Beach
Long Beach, CA
joshua.palkki@csulb.edu

Abstract

As one of America's most prominent poets, Emily Dickinson has inspired numerous composers to set her words to music. Here, we undertake the challenge of setting Dickinson's poetry with a songwriting machine, ALYSIA, that creates melodies to fit user-provided lyrics. Two new compositions are presented, and we discuss the process and co-creative features leading to their formation. A novel internal evaluation mechanism for the songwriting system and an improved generation mechanism are also discussed.

Introduction

The voice is the oldest human instrument (Kim 2003), allowing us to communicate our innermost feelings by expressing words through music. Yet, despite decades of progress in computer-aided music composition, songwriting remains primarily in the human creative domain. Can a "soulless machine" help in the human quest for songwriting? How would a machine assist without taking away from the user's self-expression? Lastly, who would benefit from such a system?

This paper explores co-creative songwriting with ALYSIA (Automated LYrical SongwrIting Application) (Ackerman and Loker 2017), a system designed to assist users in the creation of original songs by providing melodies for user-provided lyrics. In particular, machine learning models trained on human-written songs are utilized for predicting a note for each syllable in the given text. The current paper investigates ALYSIA's utility in the professional music arena, presenting co-creative features, and an evaluation mechanism for supporting collaboration between artist and machine.

ALYSIA has been endowed with two novel features designed to improve the co-creative songwriting process. First, we have added a feature that lets users explore similar melodies once a desirable melody is discovered. In addition to aiding in the discovery of melodic phrases, this feature can also be used to improve connectivity between different melodies used in a song. The second co-creative feature allows us to explore alternate rhythms for otherwise desirable melodies. This feature is particularly significant for songwriting, where rhythms enable the emphasis of certain words and syllables. This feature deepens the collaboration between human and machine, allowing the user and ALYSIA to craft the desired melody in an iterative fashion.

The ability to evaluate its own artifacts is not only an essential characteristic of a creative agent, but is also highly valuable in the co-creative songwriting process. Endowing ALYSIA with the ability to evaluate its melodies allows us to sort the melodies it creates based on their quality. As such, better melodies appear towards the top of the list. Eliminating the need to look through all of the generations not only speeds up the selection process, but also allows for a more effective search of the space of potential melodies by letting the user discover better options without substantial additional time and effort. That is, the user may ask ALYSIA to generate hundreds of melodies, and since they are presented in decreasing order based on their quality, it suffices to examine only the top few options. In addition to evaluation, substantial improvements are introduced to the generation phrase, allowing the system to better select notes for concluding musical phrases, and well as adding a new model that lets it select the octave placement of notes.

The success of the co-creative process with ALYSIA is presented through two novel compositions, with lyrics by influential American poet Emily Dickinson[1]. Human composers have found musical richness in Dickinson's work, having set her poems to music across diverse genres. But, can a machine be used to find meaningful musical phrases to capture the depth of Emily Dickinson words?[2] We discuss the co-creative process in which human composers utilize ALYSIA in an effort to create musical works based on her poetry.

Following a brief overview of previous work, we present a detailed exposition into ALYSIA's technical components, addressing both the generation and evaluation of melodies as well as the co-creative features. We then present two new songs created with ALSYIA, the first made in collaboration with Music Professor Joshua Palkki. We conclude with a discussion of future work.

[1]The first part of the title, "with bolts of melody!", is the last line of Emily Dickinson's poem *I would not paint – a picture.* (Wolff 1988)

[2]Previous work (Ackerman and Loker 2017) explored using ALYSIA's original version towards the creation of pop songs.

Previous Work

Compared with the wealth of systems for creating lyrics-free music (see, for example, an excellent overview by (Fernández and Vico 2013)), algorithmic songwriting is still in its infancy. Several songwriting systems explore the potential of autonomous algorithmic songwriting. For example, M.U. Sicus-Apparatus (Toivanen, Toivonen, and Valitutti 2013) demonstrates how the entire songwriting process can be integrated, from lyric generation to melody and accompaniment. Another interesting system, SMUG (Scirea et al. 2015), autonomously creates songs using lyrics based on academic papers. See Ackerman and Loker 2017 for a more detailed exposition of previous work on algorithmic songwriting.

Also related to our work is research on assisted music composition, which has so far primarily focused on music without lyrics ((Onisawa, Takizawa, and Unehara 2000)(Keller et al. 2006), (Hoover, Szerlip, and Stanley 2014)). A recent musical "Beyond the Fence," explored the potential of computational creativity systems to assist with the making of a full-scale stage production (Jordanous 2016). One of the most challenging aspects in the making of this musical turned out to be juxtaposition of text and melody, which stands at the heart of songwriting (Wingspan Productions 2016). This fascinating experiment demonstrated the need for new approaches when it comes to co-creating production quality songs. Recently, Ackerman and Loker 2017 introduced ALYSIA, which is the first songwriting system to learn the relationship between music and natural language. The purpose of the current work is to explore the co-creative nature of ALYSIA and investigate its utility for professional musicians, as well as enrich the system with features that facilitate the co-creative process.

Technical Description

The original version of ALYSIA relied on two models, *rhythm* and *melody*, which predict note durations and scale degrees, respectively. Given lyrics, these models are used to predict a note for each underlying syllable. The models rely on musical features, such as key and time signatures, the durations and scale degrees of the previous five notes, as well as text-based features, such as syllable type and word frequency. See Ackerman and Loker 2017 for more details.

The current version introduces several substantial changes. In particular, we incorporate an *octave* model, which predicts the octave in which each note will be positioned. Along with the remaining features, the octave model is subsequently passed to the rhythm model that predicts the durations (whole note, half note, etc.) of all generated notes. Finally, given the octave and note duration, the melody model determines the scale degrees. As in the original version, all three models rely on both text-based and musical features, and as such learn the relationship between music and natural language.

The addition of novel features enables substantial improvement over ALYSIA's previous version. Two main features have been added. The first allows the system to give special treatment to final notes in melodic lines. This lets our melody model learn how to gracefully conclude phrases, which enables the user to utilize melodies produced by our system without having to alter the concluding notes. The second feature introduces a fundamental change in melody creation. Intervals, which are the gaps between consecutive notes, are an essential musical component that is central to composition. Instead of directly encoding the knowledge of intervals, we added features that let ALYSIA learn about them. This allows our system to integrate that knowledge into melody generation. The components of the system are discussed below.

Corpus

This version of ALYSIA relies on a carefully constructed dataset, based on 30 English pop songs. Unlike MP3 files, MXL and Midi files containing lyrics are extremely rare, requiring manual construction of the data. Further, for this version, the data chosen to be used for training and testing has been vetted extensively. Instead of using an entire song as input, the system is trained on phrases. Each individual song has been broken down into 3 main categories: chorus, verse, and bridge. Furthermore, since the system is strongly dependent on the lyrics, each chorus, verse, and bridge portion were split up into musical phrases. This change lets ALYSIA train on exactly the type of data that it is expected to output: Melodies for a single line of lyrics.

This more carefully created corpus also allows us to identify the final notes in a melody, letting the system automatically select suitable phrase endings, without manually encoding rules about how melodies should be concluded. Presenting the system with properly segmented melodic lines also helps with the selection of initial notes. Further, a corpus designed in this manner prevents the system from learning from melody fragments that span multiple melodic lines.

Feature Extraction: In order to build a model, we first extract a set of features from the Music-XML (MXL) files. In addition to ALYSIA's original features (Ackerman and Loker 2017), we have added the following:

- Chorus line - If the particular input is a chorus line
- Octave - The octave range (3-6)
- Last Note: Determines whether the current note is the last note of the phrase.
- Previous notes' step difference - The interval size from one note to the next, for the first previous note down to the fifth previous note.

Model Accuracy: Using R, we trained the three models on the parsed inputs using random forest. We randomly broke down the data into a 3:1 ratio. 75% of the data was use in training and the remaining 25% for testing. Random forests are well suited for large numbers of categorical variables. Additionally, they allow non-linearity in combining features, without an explosion in the required size of the training set, in order to avoid over-fitting. Non-linearity makes random forests more powerful than linear models such as logistic regression.

A strict measure of accuracy was used, measured by iterating through the test set to find the proportion of instances on which our models made *perfect* predictions on the notes' octaves, note durations, or scale-degrees. For example, if our melody model correctly identified the scale-degree of a note then it would count towards its accuracy, whereas any other prediction would contribute to the model's error rate. The accuracy of the rhythm model on the test data is 78.76%, and for the melody model it is 65.08%.

The current corpus went through a more rigorous process and the features that have been added comment more on the general structure of music. Notably, while we observe an improved quality in the melodies, the prediction accuracy for these two models drop slightly from the previous version of ALYSIA. At times, features that appear to notably improve the subjective quality of the music led to a drop in prediction accuracy. In the interest of creating a system useful to musicians, we opt to keep such features. This brings up an interesting question: To what degree is prediction accuracy helpful in algorithmic songwriting? While it appears that it does not constitute a sufficient measure, accuracy remains an important part of any prediction-based system's success, and as such we report prediction accuracy on all three models. The accuracy of the new octave model is 90.35%, making it the most accurate of our three models. The appendix[3] shows the confusion matrices for all three models.

Song Generation: After the models are trained, we may utilize them to generate melodies for user-provided lyrics. After the lyrics are provided (given one sentence at a time), lyrics features are generated (these are features from the original version of ALYSIA, such as word-rarity, vowel strength, etc (Ackerman and Loker 2017)). For each line, we read the feature set from the lyrics and generate the octaves, followed by the rhythm and then the melody. Note that we generate one note at a time for each syllable in the lyrical phrases. We also keep track of the five previous notes, which are very important features in the model. Finally, we create a file with the generated melodies for each line, and output them in sorted order based on their quality (see below for details). The number of melodies created is specified by the users, and is otherwise unlimited.

Co-creative Functionality

In order to improve the collaboration between human artist and machine, we introduce two co-creative functionalities. Like many other creative endeavors, songwriting often consists of multiple iterations and editing. These features make it easier to explore nearby artifacts when desirable melodies are discovered.

Similarity Feature: The similarity feature allows the user to pick any generated melody and create similar ones. Similarity is achieved by inputting the last k notes of a sequence as the previous notes at the start of a new generation ($k \leq 5$ and user specified). This information allows the first few notes of the new melody to be based on a previous melody. This feature also allows separately generated phrases to connect with each other.

[3] http://www.cs.sjsu.edu/~ackerman/pub/appendix-bolts-of-melody.pdf

Alternative Rhythm Feature: This feature allows a user to identify a melody, and lets ALYSIA generate a new rhythm for the same scale degrees. This feature is helpful when a user may like a melody, but perhaps some of the note durations do not fully capture the intent of the lyrics. With this feature, the user may experiment with different variations on the same melody. This feature greatly enhances the co-creative process, allowing ALYSIA to take in feedback and provide novel insight based on that feedback.

Internal Evaluation

ALYSIA can provide any number of melodies, allowing the user to choose how many melodies to be presented with. Naturally, requesting a larger number of melodies leads to a better exploration of the space. However, sifting through dozens (or even, hundreds) of melodies is cumbersome and time consuming. Moreover, having to look through many low quality melodies before finding a few reasonable options can be frustrating.

In order to stream-line this process and allow for a more effective exploration, prior to outputting the new music file, ALYSIA performs an evaluation on each generated line. Each melody is assigned a score, based on which all melodies are subsequently sorted from highest to lowest. This feature allows the user to truly explore the space. The user may ask for a large number of melodies and only listen to the top few provided by the system.

The melodies are evaluated using Equation (1), which aims to capture two important aspects of a generated sequence, specifically the likelihood of a sequence and its entropy. Balancing the likelihood of the sequence by its entropy is essential because many pop melodies, on which we train our models, are highly repetitive. Although consecutive occurrences of the the same note (at the extreme, all having the same scale degree and duration) may work well as part of a complete composition, presenting such melodies in a co-creative system is counterproductive. Even the most amateur songwriter can set lyrics to the same repeated note, and interspersing ALYSIA's output with highly repetitive sequences only makes it more time consuming to discover interesting melodies. Adding an entropy term leads to lower scores for highly repetitive options, allowing the more varied ones to appear towards the top of ALYSIA's list of suggestions. Equation (1) has been shown to give lower rankings to sequences of repeated notes, and higher ranking to melodic, novel sequences.

Sequence Likelihood: When generating new notes for a sequence, each octave, rhythm, and scale degree chosen has a certain probability of occurring. This part of the quality equation corresponds to how likely was the model to generate a given sequence. Since the melody model uses both the octave and rhythm models, the probability of a certain scale degree incorporates the probability of the octave and rhythm models. Specifically, we attained the sequence likelihood by multiplying the probabilities of all the scale degrees in the sequence.

Entropy: The entropy is calculated by the number of unique scale-degrees occurring in a melody, normalized by

7, which is the number of (non-accidental) degrees in a scale.

Once the sequence likelihood and entropy are computed, we are ready to calculate the quality of the given melody, calculated as follows:

$$e^{-seqLikelihood^{\frac{1}{length}}} \cdot entropy \qquad (1)$$

While we may like to directly multiply the sequence likelihood by its entropy to attain the quality score, this would lead to misleading scores. The reason is that sequence likelihood is highly sensitive to the length of the sequence, as longer sequences are inherently less likely. The first factor in Equation (1) normalizes the sequence likelihood, yielding a value between 0 and 1. Note that *length* denotes the number of notes in the melody.

Figure 1: A song created with ALYSIA and Music Professor Joshua Palkki, based on an Emily Dickinson poem.

Compositions

We present two compositions based on Emily Dickinson's poetry. The first, *Hope Is The Thing With Feathers*, was co-created with ALYSIA and Music Professor Joshua Palkki. In addition to selecting and combining ALYSIA's melodies for Dickinson's poem, Dr. Palkki subsequently arranged the resulting melody into a complete musical piece, by composing a piano accompaniment that complements the melody. This points to ALYSIA's potential utility in the professional music arena.

Joshua Palkki composed with ALYSIA after most of the improvements discussed in this paper had been incorporated. Following Joshua's interactions with ALYSIA, as well as feedback from other musicians, we decided to further improve the system.

In particular, Dr. Palkki used ALYSIA with the improved corpus discussed above, and after the octave model was integrated into its architecture. The first co-creative feature, enabling the discovery of similar melodies, was also in place.

The second composition, *I Took My Power In My Hand*, was made using the most recent version of the system. This includes the second co-creative feature, which allows the user to request alternate options for the rhythm setting of ALYSIA's melodies. Based on musicians' feedback, we also radically improved the melodies themselves, which was achieved by adding features that capture interval information and phrase endings.

"Hope" is the Thing with Feathers

We present a composition co-created by ALYSIA and Music Professor Joshua Palkki, shown in Figure 1. Joshua also expanded the melody into a full composition,[4] and a recording is found at `http://mayaackerman.info/HopeIsTheThingWithFeathers.html`. We now discuss the co-creative process used by Dr. Palkki, as well as share his feedback on this co-creative experience.

Co-creative process Each line from Dickinson's poem *Hope is the Thing with Feathers* was given to ALYSIA, which was asked to provide 30 melodies for setting the line of text. Then, Dr. Palkki selected the short melodies he liked, and combined them into a single melody for the entire poem. Joshua also utilized the similarity feature, which allowed him to discover similar melodies to any of those proposed by ALYSIA. To gauge how useful ALYSIA is for assisting a professional musician, we asked Joshua to make changes to the melodies if desired. Simply put, fewer changes indicate greater utility. Most notes were left unchanged, whereas most changes that Dr. Palkki made influenced the final notes in phrases, including both their pitch and rhythm. Consequently, we decided to let ALYSIA learn how to treat final notes, by integrating a new feature that accomplishes this goal.

Expert User Experience The following was written by Joshua Palkki to describe his experience working with ALYSIA: *I very much enjoyed composing with ALYSIA. When I started working on "Hope is the Thing with Feathers," a setting of an Emily Dickinson poem, it was difficult to begin. It was my first time working with ALYSIA (or anything like it) and my first time working on an arrangement phrase-by-phrase. When I arrange a folk tune, for example, I know what the entire song/melody sounds like. Once I got started with a first few phrases that spoke to me musically, things moved quite rapidly.*

Having the computer generate melodies was exciting to me as a musician who identifies much more as an arranger than a composer. I am also a choral conductor and singer, so I worked to find melodies that honored the text (proper syllabic/word stress, for example) and also sounded pleasant and/or interesting. The advantage of working with ALYSIA was that I got to think harmonically, which is more along

[4]The full composition is found here: `http://www.cs.sjsu.edu/~ackerman/pub/hope.pdf`

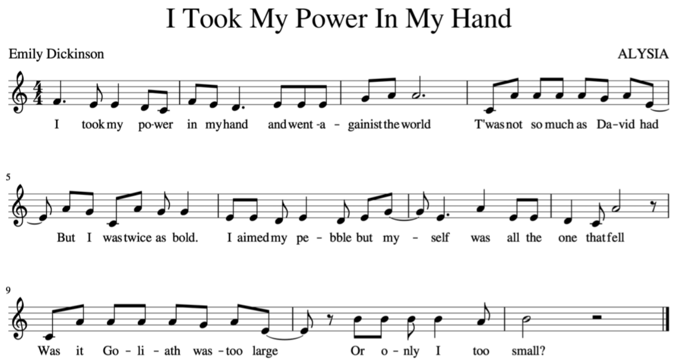

Figure 2: A song co-created with ALYSIA.

the lines of how I think as a composer/arranger - I worked to connect the harmonic "tissue" of the piece to support and enhance ALYSIAs melodies. Coming up with a melody is an incredibly difficult thing for me (and for many composers, I'd venture). Having ALYSIA provide the melodies allowed me to experience the compositional process in a new way. It has made me think that I may work very well with a collaborator (human or computer) on future projects.

We asked Joshua to describe some of the weakness of the system, so that we might improve it.

I dont think that any of the melodies that ALYSIA produced were particularly memorable on their own. But, in combination, and with an added accompaniment, I think that the result is quite beautiful. Upon reflection, this explains why the beginning of the process was so difficult. Once I had the first two snippets of melody that I found pleasing, building upon them was easier because I had a better idea of where the piece as a whole was going.

The above feedback and that of other musicians let us to further improve the melody generation through the integration of interval information. We have also asked Dr. Palkki to comment on whether ALYSIA is best described as a tool or a collaborator.

I consider ALYSIA both a tool and a collaborator. I believe that the program providing melodies helped simplify and hasten my compositional process. It was also a tool. Several times, I used the "Similar Melodies" function. I would have ALYSIA provide me 30 melodies at a time. Often I found one that I liked (often it had the best text setting) and I would ask ALYSIA to provide 30 more melodies inspired by the one I liked. It was using this function that I found some of the most interesting melodies in the piece.

I Took My Power in My Hand

The second composition was created after receiving Dr. Palkki's feedback and integrating the remaining improvements; namely, allowing the system to learn about intervals and enable special treatment of final melody notes. We've also incorporated an internal evaluation phase that lets us rank melodies based on their quality, placing the best options on top. An additional co-creative feature, letting us explore alternate rhythms for ALYSIA's melodies, was also put in.

Margareta Ackerman and Christopher Cassion used ALYSIA to created the melody found in Figure 2, and a recording is found at `http://mayaackerman.info/ITookMyPowerInMyHand.html`. Dickinson's poem, *I Took My Power in My Hand*, was given to ALYSIA one sentence at a time, and the users' role was to select amongst those options, as well as combine the short melodic phrases into a melody for the entire poem.

This song is the least edited work co-created with ALYSIA. The only changes made were to the rhythms of final note phrases, in all cases making them longer to allow the singer to gracefully conclude the phrases. We believe that the newly added features let ALYSIA create more interesting and melodic options, allowing it to take greater responsibility over the creative process. Upcoming future work will investigate how helpful the newest additions to ALYSIA are for co-creating songs with professional musicians.

References

[Ackerman and Loker 2017] Ackerman, M., and Loker, D. 2017. Algorithmic songwriting with alysia. *EvoMUSART*.

[Fernández and Vico 2013] Fernández, J. D., and Vico, F. 2013. Ai methods in algorithmic composition: A comprehensive survey. *Journal of Artificial Intelligence Research* 513–582.

[Hoover, Szerlip, and Stanley 2014] Hoover, A. K.; Szerlip, P. A.; and Stanley, K. O. 2014. Functional scaffolding for composing additional musical voices. *Computer Music Journal*.

[Jordanous 2016] Jordanous, A. 2016. Has computational creativity successfully made it beyond the fencein musical theatre? In *Proceedings of the 7th International Conference on Computational Creativity*.

[Keller et al. 2006] Keller, R. M.; Morrison, D.; Jones, S.; Thom, B.; and Wolin, A. 2006. A computational framework enhancing jazz creativity. In *Proceedings of the Third Workshop on Computational Creativity, ECAI*.

[Kim 2003] Kim, Y. E. 2003. *Singing voice analysis/synthesis*. Ph.D. Dissertation, Massachusetts Institute of Technology.

[Onisawa, Takizawa, and Unehara 2000] Onisawa, T.; Takizawa, W.; and Unehara, M. 2000. Composition of melody reflecting user's feeling. In *Industrial Electronics Society, 2000. IECON 2000. 26th Annual Confjerence of the IEEE*, volume 3, 1620–1625. IEEE.

[Productions 2016] Productions, W. 2016. Computer says show. *Two-part UK television documentary, aired on the Sky Arts channel* 25.

[Scirea et al. 2015] Scirea, M.; Barros, G. A.; Shaker, N.; and Togelius, J. 2015. Smug: Scientific music generator. In *Proceedings of the Sixth International Conference on Computational Creativity June*, 204.

[Toivanen, Toivonen, and Valitutti 2013] Toivanen, J. M.; Toivonen, H.; and Valitutti, A. 2013. Automatical composition of lyrical songs. In *The Fourth International Conference on Computational Creativity*.

[Wolff 1988] Wolff, C. G. 1988. *Emily Dickinson*. Da Capo Press.