DEVELOPING CHINESE STYLE ALGORITHMIC COMPOSITION USING MARKOV CHAINS — FROM THE CLASSICAL CHINESE-POETRY PERSPECTIVE

Jenny Ren *Phil Winsor* *Chih-Fang Huang*
Institute of Music, National Chiao Tung University

ABSTRACT

In this paper, we take advantage of Markov Chain to generate the musical structures according to the characteristic of Chinese Poetry so that with the melody, the state of mind within the Chinese Poetry could be delivered more successfully and completely. The compositional process is also a kind of sonification of Chinese Poetry. From the poesy analysis, the movement of multi-dimensional musical elements such as durations, interval relations and the overall structure can be modeled by Markov probability for stochastic algorithmic composition. We suppose that, after training, the Markov Models are used to do Stochastic Algorithmic Composition so as to turn the computer into a composer from a listener or an analyzer, where the data structure of this Markovian Algorithmic Composition is the transition matrix.[1] Through Ping-Ze Modulation, the sonification reflects the patterns in the form of Ping-Ze of each sentence in the Chinese Poetry. Eventually, the purpose of the Chinese Poetry sonification is to reinforce not only structural patterns but also image or the state of mind within the Chinese Poetry.

Keywords: Algorithmic Composition, Markov Chain, Sonification, Ping Ze.

1. INTRODUCTION

1.1. Classical Chinese Poetry

Both the classical Chinese used in Tang Dynasty and the Mandarin widely spoken today are languages based on a rich tonal system. On the subject of rules and forms of classical poetic composition, the basic requirements are focused on tonal patterns, rhyme schemes, etc. within specific length and numbers of sentences. Briefly speaking, the prosody of classical Chinese poetry, regulated by some well-defined patterns other than free intonations, is the metaphor of Chinese style composition since the classical Chinese poetry usually accompanies a song. Accordingly, the structure of classical Chinese poetry, particularly the so called Ya Yun as well as Ping Ze, could be mapped into different types of changes (i.e., increase or decrease, lengthen or shorten) in the musical elements for pitch, rhythm and dynamics.

The definitions of Ya Yun and Ping Ze are interpreted as the following: Ya Yun is literally translated as rhyming[2], which is a typical linguistic characteristic of Chinese poetry. The main purpose of Ya Yun is to make the sound harmoniously beautiful so that the recitation of the poetry sounds smooth and pleasing. Most poets put characters having similar vowels rhythmically at the end of each poetry sentence, where vowel means the end tone of a character. The other function is to make the poetry easily remembered and passed along throughout generation after generation. However, Ping Ze is literally translated as level and oblique tones[3]. It is a technical term of rhythmic poetry. Basically, the poets divide the four tones of either classical Chinese or modern Chinese into these two categories for some reasons. The level tones are flat and usually longer in sound without rising and falling tones, while the oblique tones are rising and falling tones and usually shorter in sound. These two groups of level and oblique tones are interchanged and mixed in classical Chinese poetry to produce various tonal varieties, making the poetry more interesting and less monotonous. It is the harmony of level and oblique tones which plays a significant role in classical Chinese poetry. To sum up, the use of level and oblique tones, or Ping Ze, allows classical Chinese poetry to have tones rising up and down, making the poetry rather musical in itself, and amusing to recite.

However, the pronunciation of Chinese Poetry in Tang Dynasty is close to the Southern Min, a dialect spoken in Taiwan that keeps most of the characteristics in Classical Chinese. Accordingly, the Classical Chinese Poetry is always recited in Southern Min by the Roman Pinyin. For example,

Poem Title:	*Siong-Su (Love Seed)*
Author:	*Ong5-Ui5 (Wang Wei)*
Recited Pronunciation:	Hong5-tou7 siN-lam5 kok Chhun lai5 hoat ki2 chi Goan7 kun to chhai2 hiat Chhu2 but8 choe3 siong-su

Table 1. Roman Pinyin for Classical Chinese Poetry

[1]An Nth-order Markov chain can be represented by a transition matrix, which corresponds to an N+1-dimensional probability table

[2]The term rhyme is different from the term rhythm, a sound with a regular movement or beat which is presented repeatedly.

[3]Ping means level tone while Ze means non-level or oblique tone.

1.2. Markov Chain

There are several divisions of techniques in algorithmic composition, inclusive of stochastic, rule-based flow control, grammar, chaotic and artificial intelligence. Markov Chain is one of the stochastic processes in probability theory. The Markov models have been widespread used in many other fields, like Wireless Communication and Bioinformatics.

Most of all, the principle of Markov Property is to memorize the current state. Thus, the conditional probability of future states of the process depends only on the current state, i.e. it is conditionally independent of the past states, and the path of the process, given the present state. For instance, the probability of the (N+1)th state only correlates to the current Nth state, having nothing to do with other previous states (see Fig.1). Accordingly, the Markov Process is the very process with the Markov Property, also known as Markovian.

$$\mathbf{Pr(S_{n+1} \mid S_n, S_{n-1}, \ldots, S_1, S_0) = Pr(S_{n+1} \mid S_n)}$$

Figure 1. Markov Property

Further, the Markov Chain is described as a sequence of random variables $S_1, S_2, S_3 \ldots, S_n$, with Markov Process, where each S_i is one of the possible values from a state space S. Take two lower level musical elements for example. The state space of Pitch Class is {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11} while the state space of Rhythm could be {1/1, 1/2, 1/4, 1/8, 1/16, 1/32}. The Figure 2 illustrates the process with Markov Property in Markov Chain.

$$\mathbf{S(t_0) \rightarrow S(t_1) \rightarrow S(t_2) \rightarrow \ldots \rightarrow S(t_n) \rightarrow S(t_{n+1})}$$

Figure 2. Process in Markov Chain

A Markov Chain could be represented either by a Directed Graph or a Transition Matrix. A Directed Graph consists of a set of states and a set of transitions with associated probabilities. A Transition Matrix of an N+1-dimensional probability table represents an Nth-order Markov Chain, which tells us the likelihood of an event's occurrence, given the previous N states [3, 5]. Figure 3 and Figure 4 show the 1st-order Markov Chain in terms of Directed Graph as well as Transition Matrix where C, E and G refer to Do, Mi and So, respectively.

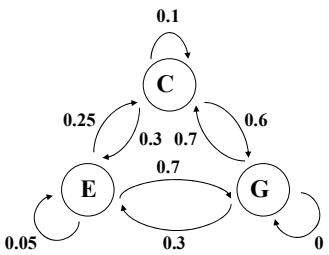

Figure 3. Directed Graph

Current \ Next	C	E	G
C	0.1	0.3	0.6
E	0.25	0.05	0.7
G	0.7	0.3	0

Figure 4. Transition Matrix

The Transition Matrix (or the stochastic matrix) P is the transition probability distribution, with (i,j)'th element of P equals to

$$p_{ij} = Pr\ (S_{n+1} = j \mid S_n = i) \qquad (1)$$

The above mentioned predicts the next 1-step state. However, when it is a stationary Markov Chain,[1] the transition matrix P is independent of the label n. Thus the n-step transition probability can be computed as the nth power of the transition matrix, P^n. In other words, $X^{(n)} = X^{(n-1)}P$, $X^{(n)} = X^{(0)}P^n$.

Besides, the higher order Markov Model remembers more additional history. The number of the order implies how many preceding states to be memorized in advance. (see Fig.5).

0th-order Markov Chain			
Current \ Next	*C*	*E*	*G*
N/A	Random	Random	Random

Independent chains have no memory, i.e. the sequence of states is a sequence of independent random variables

1st-order Markov Chain			
Current \ Next	*C*	*E*	*G*
C	20%	50%	30%
E	35%	25%	40%
G	70%	14%	16%

Transitional probability with one preceding state to be memorized

2nd-order Markov Chain			
Current \ Next	*C*	*E*	*G*
CC	15%	55%	30%
CE	20%	45%	35%
CG	60%	30%	10%
EC	35%	25%	40%
EE	49%	48%	3%
EG	60%	20%	20%
GC	5%	75%	20%
GE	0%	90%	10%
GG	70%	14%	16%

Transitional probability with two preceding states to be memorized

Figure 5. Order and Memory of Markov Chain

[1]Stationary Markov Chain, a time-homogeneous chain in itself, has one and only Transition Matrix P regardless of time. $Pr\ (S_{n+1} \mid S_n) = Pr\ (S_n \mid S_{n-1})$; Non-stationary Markov Chain, $Pr\ (S_{n+1} \mid S_n) \neq Pr\ (S_n \mid S_{n-1})$

2. MOTIVATION AND GOAL

A vast majority of the uses of Markov Chain in the algorithmic composition is to analyze and model the existing compositions. For example, some researches have already analyzed the improvisation and chord progression by means of Markov Model [1, 2, 4]. However, the Markov Models used in these studies are mainly regarded as an analyzer or a model so that the computer becomes a listener. In this study, we suppose the Markov Model do Stochastic Algorithmic Composition after training so as to turn the computer into a composer (see Fig.6). Therefore, the goal is to develop meaningful Markov Modules for musical elements so that with the advent of highly-relevant music, the emotional perception could be greatly improved during Chinese Poetry appreciation.

Figure 6. After training, the Markov Model can be used to do Stochastic Composition

3. METHODOLOGY

3.1. System Architecture

Actually, this system is the sonification process of Chinese Poetry. It models the Rhythm element via Markov Chain either in Micro Ping-Ze Modulation or in Macro Ping-Ze Modulation. The Fig.7 displays the paradigm of the compositional process.

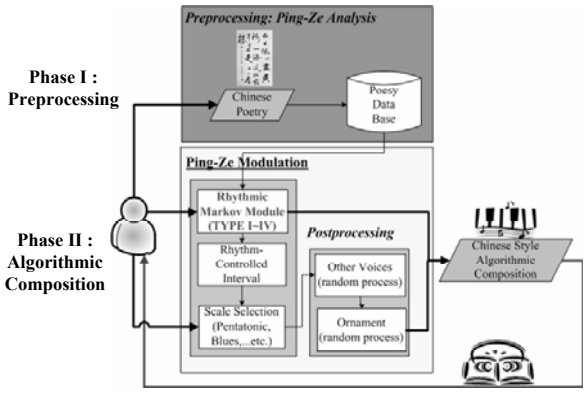

Figure 7. Chinese-Style Algorithmic Composition using Markov Chains with Ping-Ze Modulation

The implementation details are shown in Fig.8, demonstrating different user-controlled level modules:

- Module no. 0-1 & 0-2: Random Process for ornament and voices.
- Module no. 1-1 & 1-2: Ping-Ze based Markovian Process for rhythm and interval control.
- Module no. 2: User-Selected Pitch Classes, widely known as scale.

As a whole, the user input parameters are:

- Rhythmic template selection for Markovian process.
- Scale and Tonic selection.

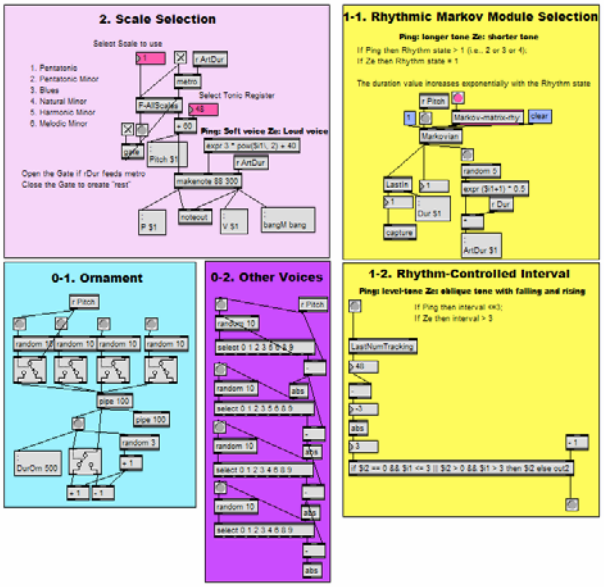

Figure 8. Max/MSP [1] Implementation

Furthermore, Fig.9 displays four basic templates for rhythmic matrices transferred from Ping Ze. Each message box has three numeric values, where the first and the second mean the current state and the next state, and the last value is the weighting of transition from current state to next state.

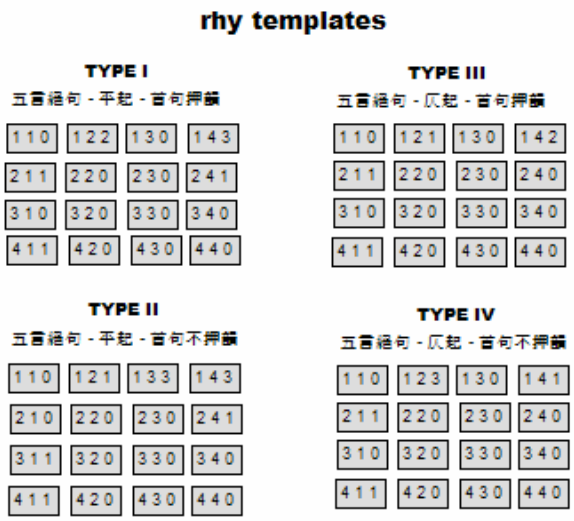

Figure 9. Transition Matrix — Four basic types of Rhythm for the Five-Character Quatrain in Chinese Poetry

[1]Max/MSP is a graphical development environment for interactive computer music and multimedia, originally written by Miller Puckette and currently developed and maintained by Cyling'74

3.2. Micro Ping-Ze Modulation — The Rhythm Markov Module

The interchanging tone and opposite tone of Ping Ze result in a rhythm effect, according to the two important characteristics in the classical Chinese poetry sentence.[1] The following shows the transformation process from Ping Ze to Rhythm. Take the Wang Wei's *Love Seed* as an example.

STEP1. Decompose phrases with Ping Ze Interchanging (where Z or 0 represents Ze; P or 1represents Ping)

Phrases	Ping Ze Structure	Pings' and Zes' aggregations
Hong5-tou7 siN-lam5 kok	P|Z|PP|Z (10110)	(10110) → *1020*
Chhun lai5 hoat ki2 chi	PP|ZZ|P (11001)	(11001) → *201*
Goan7 kun to chhai2 hiat	Z|PP|ZZ (01100)	(01100) → *020*
Chhu2 but8 choe3 siong-su	ZZZ|PP (00011)	(00011) → *02*

STEP2. Reconstruct the Rhythm Sequence

Original Poem	Ping Ze Structure	Rhythm Sequence
Hong5-tou7 siN-lam5 kok	P|Z|PP|Z 1020	*102020102002*
Chhun lai5 hoat ki2 chi	PP|ZZ|P 201	
Goan7 kun to chhai2 hiat	Z|PP|ZZ 020	
Chhu2 but8 choe3 siong-su	ZZZ|PP 02	

STEP3. Build up the Rhythmic Transition Matrix based on the Rhythm Sequence acquired previously.

	0	1	2
0	*1/6*	*1/6*	*2/3*
1	*1*	*0*	*0*
2	*3/4*	*1/4*	*0*

In step1, we first analyze the Ping Ze structure and then separate them according to the interchanging relations, where bars are allocated between P (i.e., 1) and Z (i.e., 0) in the second column. In addition, the Pings' and Zes' are aggregated to create longer or shorter sounds in each phrase. In step2, we combine and aggregate the results from step1 to reconstruct the rhythm sequence of original poem. In step3, the transitional matrix of the rhythm sequence is established. However, we normalize the number of states to a range of 4 in our implementation, as the message boxes illustrated in Fig.9, where the larger the state value, the longer the sound.

The chart below illustrates the General Ping Ze for Five-Character-Quatrain, one of the forms among Chinese Poetry in Tang Dynasty.[2] Furthermore, the definition of the four types is as the following: Type I: first-line rhyming and the second syllable being Ping; Type II: first-line without rhyming and the second syllable being Ping; Type III: first-line rhyming and the second syllable being Ze; Type IV: first-line without rhyming and the second syllable being Ze.

TYPE	Ping Ze Structure		Rhythm Sequence
TYPE I	PPZZP (rhyme) ZZZPP (rhyme) ZZPPZ PPZZP (rhyme)	201 02 020 201	*20102020201*
TYPE II	PPPZZ ZZZPP (rhyme) ZZPPZ PPZZP (rhyme)	30 02 020 201	*302020201*
TYPE III	ZZZPP (rhyme) PPZZP (rhyme) PPPZZ ZZZPP (rhyme)	02 201 30 02	*040402*
TYPE IV	ZZPPZ	020	

[1]There are two rules for interchanging tones and opposite tones. Firstly, the level (i.e., Ping) and oblique (i.e., Ze) tones in the first (original) poetry sentence are interchangeable. Secondly, the level and oblique tones in the next poetry sentence are the opposite of the previous sentence.

[2]Tang Poems are generally divided into seven forms: Five-Character-Ancient-Verse, Five-Character-Regular-Verse, Five-Character-Quatrain, Seven-Character-Ancient-Verse, Seven-Character-Regular-Verse, Seven-Character-Quatrain, Folk-Song-Styled-Verse

	PPZZP (rhyme) PPPZZ ZZZPP (rhyme)	201 30 02	*02020402*

For generalization purpose, this research basically builds up 4 Rhythmic Markov Modules for Chinese Poetry of Five-Character-Quatrain, as previously displayed in Fig.9. The Fig.10 and Fig.11 below demonstrate the implementation in Max/MSP.

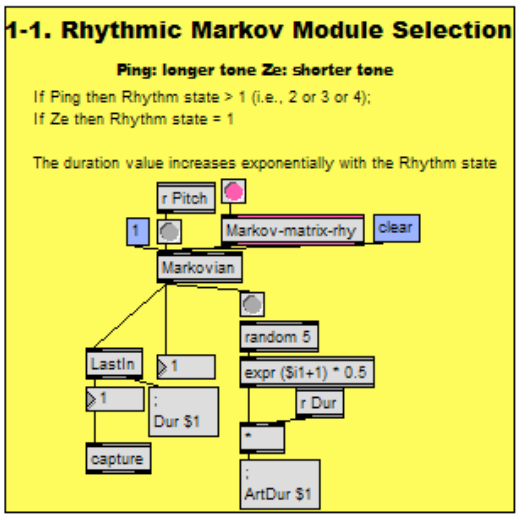

Figure 10. Rhythmic Markov Module in Max/MSP

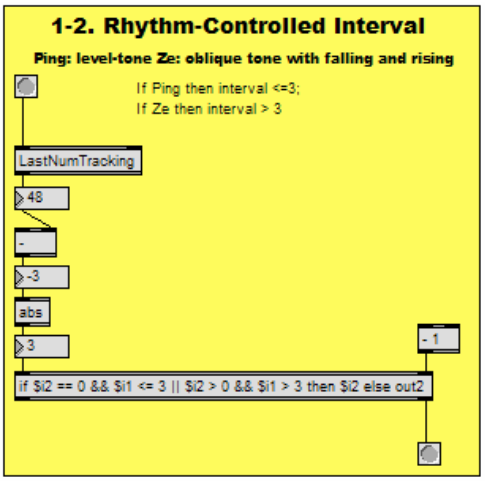

Figure 11. Ping-Ze Modulation for Rhythmic Markov Module not only controls the duration (longer or shorter tone) but also the contour of each note (level or oblique tone).

3.3. Macro Ping-Ze Modulation — The Form Markov Module

Different from Micro Ping-Ze Modulation, which aims to express the melody for some specific type (i.e., TYPE I, TYPE II, TYPE III or TYPE IV), Macro Ping-Ze Modulation contributes to a global, non-specific representation of the melody for all types on behalf of the Five-Character-Quatrain itself. That is, the former (i.e., Micro Ping-Ze Modulation) takes each interchanging tone as a rhythmic unit, while the latter (i.e., Macro Ping-Ze Modulation) regards each typical line type, or phrase structure, of tones as a common form unit. The following describes the transformation steps from Ping Ze to Form. Take the Five-Character-Quatrain which is previously illustrated above as an example.

STEP1. Categorize the penta-syllabic regulated verses (i.e., typical phrase structure of Ping Ze) into 4 common line types (A, B, C and D).

Ping Ze Structure	Rhythm	Line Type
PPZZP	201	*A*
ZZPPZ	020	*B*
PPPZZ	30	*C*
ZZZPP	02	*D*

STEP2. Label TYPE I ~ IV as sequences of Line Types.

TYPE	Ping Ze Structure		Line-Type Sequence
TYPE I	PPZZP (rhyme) ZZZPP (rhyme) ZZPPZ PPZZP (rhyme)	A D B C	*ADBC*
TYPE II	PPPZZ ZZZPP (rhyme) ZZPPZ PPZZP (rhyme)	C D B A	*CDBA*
TYPE III	ZZZPP (rhyme) PPZZP (rhyme) PPPZZ ZZZPP (rhyme)	D A C D	*DACD*
TYPE IV	ZZPPZ PPZZP (rhyme) PPPZZ ZZZPP (rhyme)	B A C D	*BACD*

STEP3. Establish the Form Transition Matrix based on the Line-Type Sequences of all TYPES.

	A	B	C	D
A	*0*	*0*	*2/3*	*1/3*
B	*2/3*	*0*	*1/3*	*0*
C	*0*	*0*	*0*	*1*
D	*1/3*	*2/3*	*0*	*0*

4. CONCLUSIONS

This study develops a public methodology to manipulate the musical elements generated from the Chinese Poetry. That is, most of the musical elements could be transformed into Markov Chain. Each Markov Module functions as a traffic sign on the way of the composition process, given the intension-oriented patterns predefined by the composer for the individual element. In this demo, the Pitch Markov Module suggests the melodic contour; the Rhythm Markov Module infers possible rhythmic patterns; the Dynamics Markov Module controls the moving intensity for a softer sound (i.e., Ping) and a louder sound (i.e., Ze). Moreover, the Form Markov Module could be extended to be a blueprint for variation in the future.

5. EVALUATION AND FUTURE WORK

Since the main idea of this study is to raise the arousal of the state of mind within original Chinese Poetry (visual stimuli) by accompanying the poems with highly-relevant music (simultaneous visual and auditory stimuli), behavioral study along with physiological measures (e.g., EDA) or neural data (e.g., EEG)[1] will be further conducted by subjects of highly skilled musicians so as to evaluate and refine the Markov Modules.

6. REFERENCES

[1] Bradley J. Clement, "Learning Harmonic Progression Using Markov Models", 1998

[2] David M. Franz, "Markov Chains as Tools for Jazz Improvisation Analysis", Blacksburg, Virginia, 1998

[3] F. Richard Moore, *Elements of Computer Music*, Prentice Hall, chapter 5, 1998

[4] Mary Farbood, Bernd Schoner, "Analysis and Synthesis of Palestrina-Style Counterpoint Using Markov Chains", *Proceedings of International Computer Music Conference*, 2001

[5] Phil Winsor, *Automated Music Composition*, University of North Texas Press, pp. 276-279, 1992

[1]EDA and EEG are abbreviations of electrodermal activity and electroencephalogram respectively.

7. APPENDIX

The following is the translation of the example poem used in this paper, a Five-Character Quatrain by Wang Wei.

相思
Poem: Love Seed

王維 (唐)
Poet: Wang Wei (Tang Dynasty)

紅豆生南國，
Red beans grow in the southern land,
春來發幾枝？
How many shoots are there in spring.
願君多采擷，
Pray gather them till full your hand.
此物最相思。
Recalling love best is this thing.

excerpted from the book "*中英對照讀唐詩宋詞*" by the author 施穎洲.