

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/275831087>

Adquisición del Conocimiento en el Proceso de Composición Musical en Base a Técnicas de Inteligencia Artificial

Conference Paper · June 2015

CITATIONS

3

READS

820

3 authors:



Efrain Astudillo

Escuela Superior Politécnica del Litoral (ESPOL)

8 PUBLICATIONS 33 CITATIONS

[SEE PROFILE](#)



Pedro Lucas

Independent Researcher

10 PUBLICATIONS 39 CITATIONS

[SEE PROFILE](#)



C. Enrique Pelaez

Escuela Superior Politécnica del Litoral (ESPOL)

51 PUBLICATIONS 1,290 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Brain Computer Interface [View project](#)

Adquisición del Conocimiento en el Proceso de Composición Musical en Base a Técnicas de Inteligencia Artificial

Efraín Astudillo¹, Pedro Lucas¹, Enrique Peláez¹

¹Escuela Superior Politécnica del Litoral, ESPOL,
Facultad de Ingeniería en Electricidad y Computación,
Centro de Tecnologías de Información,
Campus Gustavo Galindo Km 30.5 Vía Perimetral, P.O. Box 09-01-5863,
Guayaquil, Ecuador
{ejastudi, pepaluca, epelaez}@espol.edu.ec

Resumen. Generalmente, la composición musical es ejecutada por un proceso creativo en base a un amplio conocimiento de la teoría musical. Sin embargo, es posible crear piezas musicales extendiendo ese conocimiento aprendido en la academia. El propósito de este artículo es analizar y proponer representaciones abstractas, basadas en técnicas de Inteligencia Artificial, para la extracción del conocimiento en este proceso de composición e ir más allá de los tradicionales conceptos musicales. Los resultados de esta investigación contribuirán a entender la interacción de los elementos involucrados en el proceso de composición permitiendo el desarrollo de sistemas inteligentes para esta tarea y brindando soporte a los compositores en la realización de música experimental. Además, se presentan los resultados de las pruebas realizadas con una de las representaciones pertenecientes al modelo completo propuesto para el conocimiento musical, donde se generaron melodías que fueron evaluadas estéticamente y sometidas a una Prueba de Turing.

1 Introducción

Las investigaciones relacionadas con la reproducción musical basada en representaciones abstractas, utilizando técnicas de Inteligencia Artificial, se han enfocado en tres aspectos básicos: composición, improvisación e interpretación expresiva, [1] y [2]. De estos, el aspecto más explorado ha sido la composición musical, como un proceso de “armar” o ensamblar varias partes del sonido de un mismo instrumento para lograr música polifónica, o de varios sonidos distintos para crear obras musicales; como lo define el Diccionario de Música Harvard [3].

Establecer normas generales válidas para la creación musical, aplicadas a cualquier ámbito, es prácticamente imposible; por esta razón, existen estrategias de composición para ámbitos específicos, que dependiendo del entorno sugieren normas de composición, permitiendo la aplicación de técnicas definidas en la Ingeniería de Conocimiento para estructurar estas normas.

Los procesos de adquisición, validación y representación del conocimiento establecidos en la Ingeniería del Conocimiento son fundamentales en la transmisión

del conocimiento musical, análisis y su posterior reproducción; de manera que, la información sea entendible, no sólo por los ingenieros del conocimiento, que capturan el conocimiento, sino por los expertos humanos que lo transmiten, y los técnicos que lo utilizan para desarrollar sistemas basados en ese conocimiento.

Desde el punto de vista de la música, existen estándares para la transmisión de información musical. En [4] se toman en cuenta dos aspectos importantes sobre el conocimiento musical, la teoría musical y el análisis musical, cuyas interacciones a lo largo del tiempo han generado el conocimiento que hoy es impartido en la academia. Tomando en cuenta otras áreas, Christopher Alexander [5], quien realizó un estudio exhaustivo de los procesos de diseño en las construcciones arquitectónicas, establece la utilización de patrones y un lenguaje que permite la interacción con esos patrones para alcanzar lo que él denomina la *calidad sin nombre*, que en el ámbito musical sería una característica que los músicos buscan en sus creaciones y que en ciertos casos es difícil de comunicar más que en sus obras.

Las investigaciones relacionadas a composición musical [2] en su gran mayoría están enfocadas a la música occidental cuyo énfasis se centra en la música clásica y géneros estándares como el Jazz o el Blues, pero poco en la música experimental, donde según Holmes [6] existen innovadoras formas de expresar nuevas ideas para la composición produciéndose un proceso de exploración; donde más que comenzar con una partitura, se comienza con escuchar. El tipo de música experimental toma en consideración no sólo criterios elementales de armonía, melodía y ritmo, sino el tipo de sonido que se desee crear.

Las estrategias utilizadas en el ámbito de composición musical algorítmica han sido de gran interés para varias investigaciones como se detalla en [2] y [11], las cuáles son basadas en técnicas de Inteligencia Artificial. De ellas podemos hacer referencia a las principales, donde podemos encontrar: *métodos probabilísticos*, como *cadena de Markov* y *redes bayesianas* para secuencias de parámetros musicales como melodías o acordes; *sistemas basados en reglas*, que plasman la teoría musical en normas que delimitan la composición aunque con poca posibilidad de innovación; *aprendizaje de máquina*, cuya utilización se enfoca mayoritariamente a *redes neuronales* y *algoritmos evolutivos*, donde las primeras pueden representar parámetros musicales tales como notas y tiempos para generar secuencias, sin embargo pueden ser ineficientes según su aplicación; los algoritmos evolutivos en cambio pueden ser interesantes para lograr innovación en composición siempre y cuando se escoja una función de ajuste adecuada y operadores genéticos que aporten significativamente a soluciones creativas, generalmente parten de patrones musicales existentes y los modifican; *lógica difusa*, que resulta una adecuada estrategia a la hora de tratar con parámetros musicales subjetivos tales como las emociones transmitidas por la música; *programación por restricción*, donde las restricciones para atributos musicales tales como el ritmo pueden llevar a resultados aceptables y eficientes; *gramáticas generativas*, las cuales estructuran parámetros musicales para dar sentido a la composición; y finalmente *híbridos entre la estrategias* que permiten establecer resultados con mayor coherencia por medio de una composición modularizada.

Es importante destacar que algunas de estas estrategias utilizan la teoría lingüística para representar reglas de composición musical, e identificar patrones; tales como, frases melódicas, transiciones, arreglos de armonías, entre otros [2], de esta forma se brinda variabilidad a la salida del algoritmo de composición, y meta-conocimiento

como criterio de evaluación para “guiar la composición” más que la “composición” en sí.

El modelo para la adquisición del conocimiento, desarrollado en esta investigación, permite crear una base de conocimientos que toma en consideración la experiencia humana de dos compositores músicos entrevistados, relacionada a la composición de música experimental y las técnicas de inteligencia artificial aplicadas al proceso creativo de los artistas, bajo el mismo esquema que utilizarían cuando crean este tipo de música.

En la sección 2 de este artículo se muestra la metodología utilizada para extraer y representar el conocimiento de los expertos y utilizarlo en un experimento de generación de melodías musicales para una parte del modelo de adquisición generado por y para ese conocimiento; la sección 3 presenta la arquitectura de adquisición y la descripción de sus partes; en la sección 4 se presentan los resultados de la experimentación basadas en criterios estéticos; y en la sección 5 se detallan las conclusiones y futuros trabajos con respecto a esta área.

2 Metodología

Para la adquisición del conocimiento humano en el proceso de composición musical se aplicó una entrevista no estructurada a dos músicos, siguiendo los lineamientos de Kendal & Creen [7]. Los músicos entrevistados son compositores con formación académica en música y practicantes de la música experimental, además fueron familiarizados con el proceso de adquisición a utilizar. Éste proceso de adquisición será detallado en profundidad en la sección 3.

Los lineamientos extraídos de la entrevista realizada permitieron adquirir el conocimiento general del proceso creativo. Siendo este primer componente el *conocimiento explícito* representado a través del lenguaje natural, y en base a estos lineamientos se diseñó una arquitectura tecnológica de adquisición del conocimiento para capturar un segundo componente que es el *conocimiento tácito*, que por su particularidad no puede ser articulado o transmitido fácilmente, el cual está implícitamente representado en las melodías musicales.

Posteriormente, se desarrolló uno de los componentes de la arquitectura tecnológica para representar este conocimiento, cuyo proceso consistió en crear las matrices de transiciones de una Cadena de Markov de primer orden, que constituye la primera representación del conocimiento extraído de los músicos, basadas en las melodías compuestas por ellos. Este proceso dio lugar a la extracción del conocimiento tácito dado por las grabaciones de secuencias de notas hechas por los músicos para melodías, las cuales fueron almacenadas en archivos de audio MIDI.

Las nuevas melodías generadas a partir del conocimiento extraído fueron producidas mediante un algoritmo de generación de secuencias descrito en la sección 3.4.

Un grupo de melodías generadas artificialmente y otro grupo de ellas compuesta por los músicos entrevistados fueron entregadas a otros músico con formación académica con el objetivo de calificarlas estéticamente y luego clasificarlas, según su criterio, en dos grupos; las producidas por un humano y producidas por una máquina.

Con el objetivo de evitar un sesgo en los resultados, inicialmente para las calificaciones no se les comunicó que dos entes distintos produjeron las melodías, luego se les mencionó este hecho para que pudieran clasificarlas.

Finalmente, se estableció un análisis de la experimentación y sus implicaciones en el avance de una investigación que busque probar toda la arquitectura propuesta.

3 Arquitectura para el Sistema de Adquisición y Representación del Conocimiento

Para generar una arquitectura capaz de adquirir y representar el conocimiento implícito que poseen los compositores de música, se ha tomado en consideración las descripciones e idiosincrasias de los entrevistados, y se han estructurado los patrones específicos que les permitan componer con su estilo. La música experimental de los entrevistados es de carácter occidental, por lo tanto, su representación estará basada en el sistema de doce notas con frecuencias específicas, así:

C – C# – D – D# – E – F – F# – G – G# – A – A# – B

Estas notas constituyen los elementos atómicos básicos para la representación de música occidental. Estas notas pertenecen a la escala musical cromática, y son la base para generar otras escalas y acordes, la que nos permite producir representaciones generales que pueden ser almacenadas en nuestra base de conocimientos tomando en cuenta las que han sido utilizadas en [8] y [9]; por ejemplo:

Si asignamos un orden a la escala cromática anterior, tendríamos:

C	C#	D	D#	E	F	F#	G	G#	A	A#	B
0	1	2	3	4	5	6	7	8	9	10	11

Existen varias combinaciones estándares de escalas y acordes; estas combinaciones pueden ser patrones que son posibles representar en base a la nota clave que constituyen; por ejemplo:

La escala C Major está compuesta por las notas: C, D, E, F, G, A, G. En esta escala la nota clave es C; si representamos C como X, y tomamos en cuenta el orden asignado a la escala cromática anterior, podemos representar la escala X Major de la siguiente como **X Major**: (X, X + 2, X + 4, X + 5, X + 7, X + 9, X + 11).

En general la representación de un patrón de melodía bajo una clave musical X sería:

Patrón de melodía X: (X, X + a, X + b, X + c, ... X + n); tal que $a, b, c, n \in \mathbb{Z}$. (1)

Extrapolando esta representación para acordes y considerando que son notas ejecutadas simultáneamente, tenemos:

$$\text{Patrón de acorde X: } \begin{pmatrix} X \\ X + a \\ X + b \\ X + c \\ \dots \\ X + n \end{pmatrix}, \text{ tal que } a, b, c, n \in \mathbb{Z}. \quad (2)$$

Para la representación del tiempo de ejecución E, tenemos:

$$E : (\text{nota}, te, tp) . \quad (3)$$

Donde *nota* es la nota musical ejecutada, *te* es el tiempo de ejecución (por ejemplo, el tiempo que se mantiene una tecla presionada), y *tp* es el tiempo del silencio antes de ejecutar la siguiente nota. Los tiempos son relativos al tempo general de una pieza y se expresan en BPM (beats por minuto).

Estos elementos atómicos básicos son parte fundamental en las salidas de la siguiente arquitectura del sistema de adquisición y representación del conocimiento propuesto:

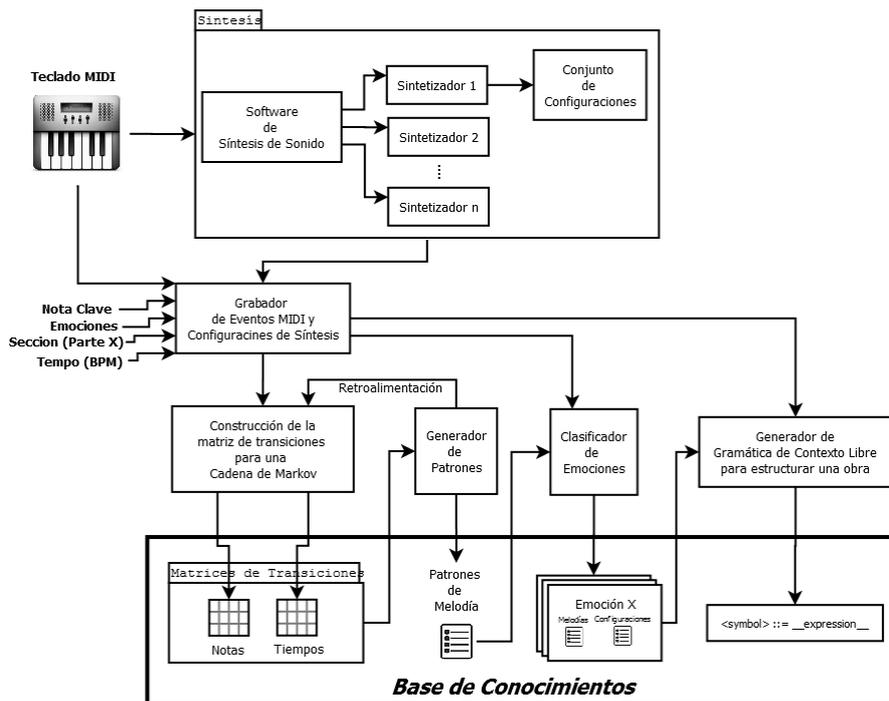


Fig. 1. Diagrama de bloques para la arquitectura del sistema de adquisición del conocimiento

A continuación se describen los componentes de esta arquitectura tomando como ejemplo una obra de uno de los compositores con gran influencia en los músicos

entrevistados, consideraremos *la Toccata y fuga en re menor, BWV 565 de Johann Sebastian Bach* que es muy conocida mundialmente y controversial por sus componentes únicos para la época.

3.1 Entradas

Tal como se muestra en la figura 1, los datos de entrada son:

- La información de la ejecución de la composición generada por los músicos; es decir, una secuencia de eventos MIDI de un teclado musical que consta de tres elementos: nota musical que se está ejecutando (0 a 127), velocidad (0 a 100) y duración (BPM beats por minutos).
- La nota clave (*key*) que rige la interpretación del músico, relativa a los patrones que se van a extraer.
- La emoción que expresa el sentimiento de lo que se está interpretando en ese momento (alegría, nostalgia, ira, etc.).
- La sección o tramo de la interpretación, tal como fue asignado por el autor a su composición; es decir, la parte A, B, C, etc., de la composición, cada una de las cuales podría tener su conjunto particular de emociones. Y,
- El tempo o la restricción de tiempo dada en BPM (beats por minuto), guiado por un metrónomo de sincronización entre la interpretación y la grabación que el sistema realiza.

Considerando la obra de Bach, podemos estructurar un conjunto hipotético de entradas que recibiría el sistema de adquisición y representación basado en el criterio de los músicos de esta manera:

Teclado MIDI: A, G, A, G, A, E, D, C#, D..... (Cada una debe ser acompañada por su velocidad y tiempo)

Nota Clave: D menor

Tempo: 80 BPM

Emociones y Partes:

Tabla 1. Partes y emociones de la Toccata y fuga en re menor, BWV 565 de Johann Sebastian Bach

Intervalo de tiempo (mm:ss)	Parte	Emociones
[0:00, 00:24)	A	Ira
...
[02:48, 04:05)	F	Melancolía, Nostalgia, tristeza
...
[08:11, 09:10)	J	Sensación de libertad

3.2 Síntesis

En este módulo se producen los sonidos de las composiciones por medio de sintetizadores digitales de los cuáles se almacenan sus configuraciones con el fin de enlazarlas con los datos recolectados en el módulo de grabación.

En el caso de la obra de Bach es una composición barroca que generalmente es interpretada por un órgano, por lo tanto una configuración de síntesis para esta obra conllevaría a registrar la configuración de elementos de un sintetizador tales como osciladores, filtros y envolventes de tal manera que simulen el modelamiento físico del órgano.

3.3 Grabador de Eventos MIDI y Configuraciones de Síntesis

Este módulo recolecta las entradas y las procesa de acuerdo al requerimiento de los otros módulos del sistema; también, almacena la interpretación combinada con el resto de entradas, de tal manera que puedan reproducirse después. La utilización de aplicaciones de grabación MIDI reduce la complejidad de la implementación.

3.4 Representación del Conocimiento a través de una Matriz de Transiciones con Cadenas de Markov

Las Cadenas de Markov son un proceso estocástico secuencial de estados o eventos dependientes. Donde la transición de un estado al otro depende sólo del estado actual y en su probabilidad de ocurrencia. Los dos componentes principales de las cadenas de Markov son: el conjunto de estados y las transiciones entre ellos [11].

Debido a esta característica, las cadenas de Markov cuentan con un gran potencial para la representación de frases melódicas donde los estados de dicha cadena pueden ser dados por valores de notas MIDI, frecuencias, tiempos, o duraciones, entre otros. En nuestro trabajo utilizamos los valores de las notas MIDI (enteros positivos).

Las cadenas de Markov pueden ser representadas por *matrices de transiciones*; una matriz de este tipo posee las siguientes características [10]:

1. Es una matriz cuadrada. Si se tiene un conjunto de estados S con n estados. La matriz será de $n \times n$.
2. Todos los valores de la matriz están entre 0 y 1.
3. La suma de los valores en cualquier fila debe ser igual a 1.

Para construir estas matrices en cada celda se calcula la probabilidad de ir de un estado j , dado que el estado actual es i , siendo i y j coordenadas de la matriz para filas y columnas respectivamente. Esta probabilidad está dada por la relación entre el número de ocurrencias en el conjunto de elementos para una secuencia ij y la suma de las ocurrencias de las secuencias $i0$ hasta ij ; es decir, la suma de las ocurrencias en la fila i , de tal manera que se cumpla que la suma de esa fila sea igual a 1. Este proceso vendría a ser la estrategia de entrenamiento para una cadena de Markov.

El pseudocódigo descrito a continuación es un algoritmo para generar secuencias a partir de una matriz de transiciones previamente entrenada, iniciando a partir de un estado arbitrario [13].

Pseudocódigo para generar el siguiente elemento en una secuencia usando la matriz de transiciones [13].

```

float transition_matrix[N][N]; //matriz de probabilidades de la
tabla 3
char estados[N]; //representación del conjunto de estados S
function Generate_next_element(int current_element):
float valor;
float limite_superior = 0.0F;
begin
    valor = OBTENER_NUMERO_RANDOM // entre 0.0 y 1.0
    for_each index in [0..N]:
        limite_superior += transi-
tion_matrix[current_element][index];
        if ( valor <= limite_superior )
            return index; // índice del estado siguiente
        end
    end
end

```

Cadenas de Markov como Método de Representación para la Arquitectura

Propuesta. En Nierhaus [11] se propone la representación de cadenas de Markov para composición musical. En base a ello y al conocimiento adquirido de los músicos entrevistados, en este módulo se generan dos matrices, una para las transiciones entre notas musicales; y, otra para las transiciones entre tiempos y duraciones, representados por (te, tp) . Esta representación nos permite a futuro generar composiciones más ricas, combinando patrones de tiempo con patrones de secuencias de notas que no necesariamente utilizaron esos patrones de tiempo y generar así conocimiento nuevo, que no fue adquirido directamente de los expertos.

Por ejemplo, las matrices de transiciones de notas y duraciones para la Toccata y Fuga de Bach presentan una particularidad en sus melodías, donde algunas de ellas no se encuentran dentro de la clave D menor, y las duraciones se pueden encontrar en su mayoría representadas por semicorcheas (0.375 segundos si BPM es 80) [12]. A continuación se presenta la matriz de transiciones de notas para esta obra, construida con el algoritmo para las cadenas de Markov.

Tabla 2. Matriz de transiciones de notas de la Toccata y Fuga en re menor, BWV 565 de Johann Sebastian Bach

	C	C#	D	D#	E	F	F#	G	G#	A	A#	B
C	0.06	0.01	0.19	0.04	0.07	0.09	0.02	0.08	0.00	0.11	0.27	0.05
C#	0.02	0.04	0.36	0.00	0.24	0.02	0.00	0.02	0.00	0.15	0.13	0.03
D	0.11	0.08	0.21	0.05	0.12	0.08	0.02	0.10	0.01	0.15	0.06	0.01
D#	0.15	0.02	0.42	0.02	0.00	0.23	0.02	0.08	0.00	0.02	0.02	0.02
E	0.05	0.13	0.19	0.00	0.05	0.23	0.01	0.16	0.00	0.15	0.02	0.00
F	0.04	0.02	0.17	0.05	0.19	0.04	0.00	0.23	0.01	0.17	0.06	0.01
F#	0.02	0.01	0.22	0.09	0.03	0.04	0.01	0.35	0.01	0.18	0.05	0.00
G	0.08	0.01	0.09	0.03	0.14	0.17	0.04	0.05	0.01	0.23	0.11	0.04
G#	0.00	0.00	0.17	0.00	0.08	0.06	0.04	0.17	0.12	0.25	0.00	0.12
A	0.03	0.08	0.17	0.00	0.08	0.12	0.02	0.19	0.00	0.11	0.17	0.04
A#	0.11	0.10	0.10	0.02	0.05	0.02	0.01	0.20	0.01	0.34	0.03	0.01
B	0.24	0.22	0.09	0.01	0.01	0.04	0.00	0.17	0.01	0.17	0.00	0.04

Se puede apreciar en la matriz anterior que apenas 18 transiciones de las 144 posibles tienen probabilidad cero lo cual nos da un indicio de que las líneas melódicas no se restringen de una manera estricta a D menor.

3.5 Generador de Patrones

En este módulo de la arquitectura propuesta, se crean los patrones melódicos usando un corpus de pequeños segmentos de melodías musicales. Cada segmento es una limitada secuencia de notas con sus respectivas duraciones generadas por las cadenas de Markov y deben cumplir con ciertas reglas musicales usadas en el proceso de composición. Adicionalmente, se agregan otros segmentos a partir de pequeñas melodías musicales que han sido grabadas a través del tiempo por los músicos, de igual manera deben ser previamente analizadas y procesadas para que cumplan con las reglas musicales mencionadas anteriormente. Estos patrones son representados de la siguiente manera:

Patrón de Melodía M : $(S_1, S_2, S_3, \dots, S_n)$

Donde los elementos de M son segmentos de melodías, y cada segmento S_i , contiene: $(nota_i, te_i, tp_i)$

El proceso para generar estos patrones consiste en combinar N diferentes segmentos del corpus, formando una secuencia concatenada de estos segmentos. Debido a que el corpus puede ser vasto, debemos realizar una búsqueda informada entre las diferentes soluciones. Los resultados de esta búsqueda deben ser las características fundamentales y estructurales de una composición musical que debe cumplir cada patrón y tener coherencia perceptiva en el oyente.

En el proceso de composición existe una retroalimentación al módulo de construcción de matrices para facilitar el proceso iterativo que ocurre durante la composición, y así generar más patrones que son el resultado del autoaprendizaje del sistema. Algunos de esos patrones no necesariamente son congruentes con las emociones que se desean transmitir, de manera que este modelo concibe un módulo clasificador de emociones, para relacionar estos patrones.

Por ejemplo, en la obra de Bach se generan secuencias de notas con largas duraciones y esa secuencia es similar a las notas musicales de una parte que evocaba esperanza, pero con duraciones cortas, posiblemente, el resultado emocional sería más sublime apuntando a sensaciones positivas.

3.6 Clasificador de Emociones a través de Lógica Difusa

Una de las entradas al sistema son las etiquetas de emociones que el músico coloca a sus composiciones. Este módulo permite asignar esas etiquetas a los patrones generados por el sistema, dependiendo de su nivel de relación con patrones ya etiquetados; es decir, si son patrones nuevos provenientes del generador de patrones discutido en la sección anterior, y deseamos saber si pertenecen a la emoción

“alegría”, basta con comparar con los patrones almacenados de alegría, y determinar su grado de cercanía a través de *lógica difusa*; esta comparación también nos permite determinar el grado de membresía que podría tener con respecto a otras emociones. Un proceso similar se realizaría para los sonidos producidos por el módulo de síntesis cuyo elemento no serían patrones de melodías, más bien configuraciones de síntesis.

Considerando que es posible representar a los elementos que participan en la clasificación como secuencias que podrían ser almacenadas en vectores, el grado de cercanía puede ser determinado por cálculo de distancias tales como la Euclidiana, de Coseno, Manhattan, entre otras.

Este módulo permite la representación del “rango de emociones” que los músicos mencionaron en la entrevista y que desean plasmar en sus composiciones; emociones que en este modelo son representadas a través de funciones de membresía para los patrones.

Por ejemplo, la figura 2 muestra las funciones de membresía correspondientes a patrones melódicos con respecto al ejemplo de emociones identificadas en la tabla 1 para la obra de Bach. Nótese que en el eje X (patrones melódicos), los valores P_i serán discretos y representados por arreglos de patrones, pues se debe considerar la posibilidad de que existan elementos con el mismo grado de membresía.

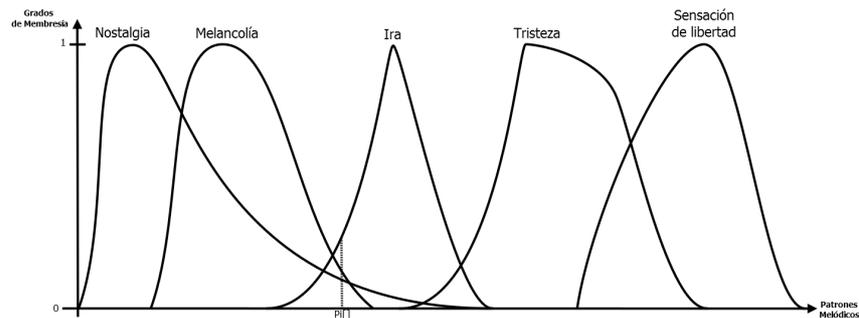


Fig. 2. Funciones de Membresía para patrones melódicos con respecto a emociones

3.7 Generador de Gramática de Contexto Libre para Estructurar una Obra

A través de este módulo se logra estructurar la obra como un todo a partir de la definición de las secciones (partes), los patrones de melodías y las configuraciones del sintetizador. El producto de este módulo es el conjunto de reglas que representan la gramática de contexto libre, a través de la cual se pueden generar las estructuras de la obra.

Para obtener esta gramática se combinan dos estrategias: Obtener reglas de los expertos, que permitan clasificar las partes, y a través de un mecanismo de *inferencia de gramática* obtener las reglas desde las composiciones.

Por ejemplo, para una gramática de la estructura de Johann Sebastian Bach, considerando el criterio de los músicos y patrones genéricos sería:

```

<Parte A> ::= 'Patrón 1' | 'Patrón 2'....
<Parte C> ::= 'Patrón 9' | 'Patrón 3' | 'Patrón 10' |....
<Parte F> ::= <Parte A> | 'Patrón 5' | <Parte B> <Parte H>
<Parte C> |....
<Parte I> ::= <Parte D> | <Parte F> 'Patrón 6' | ....
.....

```

4 Resultados

En esta investigación y para validar el modelo, se pidió a cinco músicos compositores con formación académica evaluar melodías generadas por el sistema y otras producidas por los músicos entrevistados, calificándolas en un rango de 0 a 10, siendo el valor más alto la mejor calidad de la melodía según el criterio de cada músico evaluador.

El banco de melodías compuestas por los músicos experimentales fueron 30; 15 de ellas fueron usadas para el entrenamiento del sistema, y las otras 15 fueron utilizadas para evaluar el sistema. Las 15 melodías de entrenamiento tuvieron un total de 845 notas con un promedio de 56.33 notas para cada secuencia de melodía; lo mismo fue aplicado para las duraciones de cada nota, lo que en conjunto alimentó las matrices de transiciones. También, se generaron 30 melodías por el sistema, luego de ser entrenado, y a través de la estrategia de generación de secuencias para las cadenas de Markov. En total se entregaron 45 melodías a los músicos evaluadores, ordenadas al azar.

Para establecer el tipo de prueba estadística a aplicar, fue necesario verificar la normalidad de las muestras. En la tabla 3 se presentan los resultados obtenidos por el test de normalidad de Shapiro-Wilk.

Tabla 3. Resultados de la prueba de Shapiro-Wilk

Compositor	Valor p	W
Humano	0.003841	0.9479
Máquina	3.895e-07	0.924

Podemos observar que los datos no siguen una distribución normal; por lo tanto, se optó por la prueba Wilcoxon signed-rank [12]. Wilcoxon signed-rank es una prueba no paramétrica que no asume la normalidad de los datos. Es usada para comparar dos grupos que vengan de los mismos participantes, con la finalidad de observar el cambio en los sujetos de un tiempo a otro, o cuando los individuos son sometidos a diferentes condiciones y analizar los efectos causados en los individuos. Para la prueba se asume que los dos tipos de melodías siguen el mismo estilo de los músicos que las generaron.

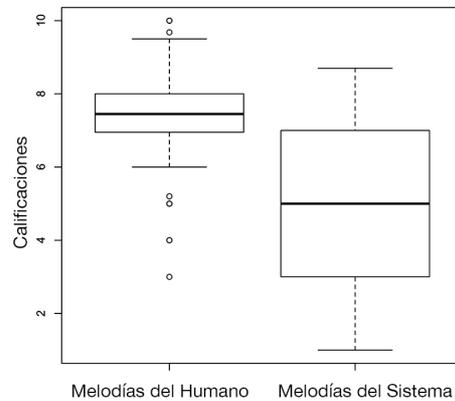


Fig. 3. Diagrama de cajas para los grupos de melodías evaluados

Tabla 4. Estadística descriptiva de los dos grupos

Compositor	N	Mín.	Max.	Media	Mediana	Desv. Stand.
Humano	75	3.00	10.00	7.343	7.450	1.457167
Máquina	150	1.00	8.700	4.902	5.00	2.109666

En la figura 3 y la tabla 4 se muestra un diagrama de cajas y la estadística descriptiva entre los dos grupos de melodías respectivamente, se puede observar que las medianas difieren significativamente, lo cual fue comprobado por la prueba de Wilcoxon signed-rank, como se muestra en la tabla 5, de manera que podemos verificar que las melodías generadas por el sistema no siguen los mismos patrones que se le administró para el entrenamiento con el 95% de confiabilidad.

Tabla 5. Valores de la prueba Wilcoxon signed-rank

Valor p	W
4.733e-16	9355

La hipótesis nula en la prueba Wilcoxon signed-rank define que las diferencias de medianas entre pares de observaciones es igual a cero, por consiguiente el valor **p**, obtenido en la investigación muestra que existe una diferencia estadística significativa de las medianas de estos dos grupos de melodías.

Con respecto a la prueba de Turing realizada, se tomaron en cuenta los aciertos con respecto a la clasificación en cada grupo (humano o sistema) como lo muestra la figura 4.

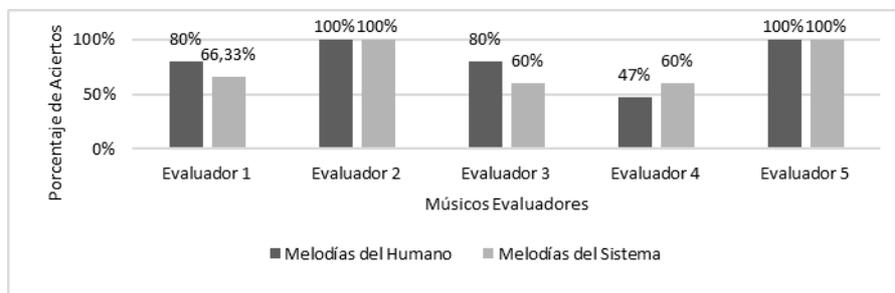


Fig. 4. Porcentajes de aciertos para la prueba de Turing

Nótese que dos de los evaluadores (el 2 y el 5) lograron clasificar el material correctamente, el resto posee un porcentaje de fallo que indica que hubo melodías que no pudieron ser distinguidas para el grupo al que pertenecían. Además, los comentarios con respecto las pruebas reportan que las melodías que ellos creían que eran generadas por humanos tenían estructura y coherencia de inicio a fin, mientras que las demás se sentían más aleatorias y con una estructura difusa, adicionalmente mencionaron que la calidad pudiera aumentar si existiese una base armónica que soportase cada melodía.

5 Conclusiones

El diseño de la arquitectura de adquisición y representación, propuesta en este trabajo, es un punto de partida para poder estructurar el conocimiento musical intrínseco en un artista, es flexible con respecto a la aplicación de técnicas de Inteligencia Artificial, como fue en este caso la representación de cadenas de Markov. Además, la estrategia de generación de melodías aplicada a las matrices de transiciones promete ser un enfoque adecuado si se toman en cuenta otros factores que mejoren el procedimiento, como por ejemplo una base armónica sobre la cuál sea reproducida una melodía en particular.

Es importante destacar el beneficio incremental que se podría obtener estéticamente en composiciones basadas en esta arquitectura, hemos demostrado que este primer paso en el desarrollo de un sistema de composición, a pesar de no ser suficiente para lograr exhaustivamente la expresividad que un compositor humano puede generar al componer, produjo varias melodías, del conjunto generado, que algunos de los músicos evaluadores no pudieron distinguir si pertenecían a un humano o una máquina.

Las estrategias de representación del conocimiento para músicos experimentales artificiales, deben tomar en cuenta el proceso creativo tanto en lo explícito como lo implícito; es decir, argumentar las representaciones en base a la identificación de sus ideologías, influencias, actividades e incluso equipos como los sintetizadores, lo cual sería lo explícito, y en lo implícito extraer ese conocimiento directamente de sus creaciones.

La implementación de sistemas de composición e interpretación, que asistan a músicos en su proceso creativo, debe considerar los elementos de la obra que ellos quisieran mejorar estéticamente, esto implica en primera instancia un entendimiento del proceso de composición; como el realizado en esta investigación, con el objetivo de generar recomendaciones por parte del sistema; sin embargo, existen límites impuestos por los artistas, como el no aceptar que la máquina sea un músico más, y por tanto el enfoque de este tipo de sistemas debe ser de colaboración, con el objetivo de cortar camino para llegar a componer obras intensamente emocionales y condensadas, tal como lo mencionaron los músicos entrevistados.

A futuro, una posibilidad es considerar la utilización de algoritmos genéticos para la generación de patrones melódicos; así como explorar la calidad de las composiciones generadas por el módulo de clasificación de emociones, tomando en cuenta los distintos tipos de métodos de cálculo de distancia y funciones de membresía para la asociación de patrones melódicos y emociones.

Referencias

1. Lopez de Mantaras, R.: Making Music with AI: Some examples. In Proceedings of the conference on Rob Milne: A Tribute to a Pioneering AI Scientist, Entrepreneur and Mountaineer. IOS Press (2006) 90–100
2. Fernández, J., Vico, F.: AI methods in algorithmic composition: a comprehensive survey. *Journal of Artificial Intelligence Research*, Vol 48 (2013) 513-582
3. Willi, A.: *Harvard Dictionary of Music*. 2nd edn. The Belknap Press of Harvard University Press. Cambridge, Massachusetts (1974).
4. Lilienfeld, R.: An Introduction to Music. *Music Educators Journal*, Vol. 48 (1962).
5. Alexander, C.: *The Timeless Way of Building*. New York Oxford University Press (1979)
6. Holmes, T.: *Electronic and Experimental Music: Pioneers in Technology and Composition*. 3rd edn. London and New York: Routledge (2008)
7. Kendal, S.L., K, Creen, K.: *An Introduction to Knowledge Engineering*. Springer-Verlag London (2007)
8. Ulrich, J.: The analysis and synthesis of jazz by computer. Fifth International Joint Conference on Artificial Inteligence (1977)
9. Ganesh, R.S., Palaniappan, C.T., Ramakrishnan, M.S., Devanathan, R.: Knowledge Engineering of Creative Musical Expressions Using Carnatic Music Ideology. International Florida Artificial Intelligence Research Society Conference (2004)
10. Greenwell, N.R., Ritchey, P.N., Margaret, L.: *Calculus with applications for the life sciences*. Pearson Education (2002)
11. Nierhaus, G.: *Algorithmic Composition: Paradigms of Automatic Music Generation*. Springer-Verlag/Wien (2009)
12. Toccata and Fugue in D minor, BWV 565, http://en.wikipedia.org/wiki/Toccata_and_Fugue_in_D_minor,_BWV_565
13. Markov chain implementation in C++ using Eigen, <http://www.codeproject.com/Articles/808292/Markov-chain-implementation-in-Cplusplus-using-Eig>