

UNIVERSIDAD DE SALAMANCA

TESIS DOCTORAL

---

# Sociedades Humano-Agente: Un Caso de Estudio en Creatividad Musical

---

*Autor:*

María NAVARRO

CÁCERES

*Director:*

Dr. D. Juan Manuel CORCHADO RODRÍGUEZ

*Co-director:*

Dr. D. Marcelo FREITAS CAETANO



UNIVERSIDAD  
DE SALAMANCA

*La presente tesis ha sido presentada en cumplimiento de los requisitos  
para la obtención del grado Doctor*

*en el*

Departamento de Informática y Automática

14 de Junio de 2017



# Declaración de Autoría

D<sup>a</sup> María NAVARRO CÁCERES, presenta el proyecto de tesis titulado “Sociedades Humano-Agente: Un Caso de Estudio en Creatividad Musical” para optar al Grado de Doctor en Informática y Automática por la Universidad de Salamanca, y declaro que ha sido realizado bajo la dirección del profesor Dr. Juan Manuel Corchado Rodríguez, profesor catedrático del Departamento de Informática y Automática, y el investigador Dr. Marcelo F. Caetano, investigador post-doctoral en el Instituto de Engenharia de Sistemas e Computadores (INESC TEC) de la Universidad de Porto, Portugal.

En Salamanca, a 14 de Junio de 2017

El doctorando:

Fdo: María NAVARRO CÁCERES

Los directores:

Dr. D. Juan Manuel CORCHADO RODRÍGUEZ

Dr. D. Marcelo FREITAS CAETANO



*“Hermoso es lo que vemos. Más hermoso es lo que sabemos. Pero mucho más hermoso es lo que no conocemos.”*

Niels Steensen



## *Resumen*

Los contextos sociales son esenciales para la toma de decisiones humanas, proporcionando la base para las inferencias y la planificación de actividades. Computacionalmente, esta información social se incorpora al concepto de sociedades humano-agente, y se aplica en campos de investigación muy diversos, como la Medicina, la Seguridad Cibernética o la Economía. Las sociedades humano-agente permiten la colaboración entre una organización social humana y otra virtual para intercambiar información y llegar a conseguir una serie de metas, que de forma individual serían imposibles de conseguir. En este caso, se pretende construir una sociedad humano-agente en un contexto musical, donde los seres humanos y las máquinas se coordinan para generar música a partir de una entrada proporcionada por los usuarios humanos, que puede consistir en una configuración tonal para generar acordes, o los colores de una imagen digital para generar una melodía. La sociedad se implementa como una Organización Virtual (VO), que contiene dos módulos de composición. El primero de ellos interacciona con el usuario para generar una progresión de acordes. El segundo de ellos es capaz de utilizar la opinión social para transformar los colores de una imagen digital en una melodía que sea aceptable para la comunidad que participa en esta sociedad. Para evaluar el enfoque actual, una pequeña sociedad de expertos califica el sonido producido siguiendo criterios de consonancia y musicalidad.



# *Abstract*

Humans need social contexts in order to make decisions; they are a basis that makes inferences and planning activities possible. Computationally, social information is incorporated to the concept of human-agent societies, and it is applied in very diverse fields of research, such as Medicine, Cyber-Security and Economy etc. Human-agent societies enable collaboration between human and virtual social organizations for the exchange of information and the achievement of a series of goals, which individually would be impossible to achieve. In the present work, we intend to build a human-agent society in a musical context, where humans and machines are coordinated to generate music. This is done when an input is provided by human users. An input can consist of either a tonal configuration which is used to generate chords or from the colors of a digital image, used to generate a melody. The society is implemented as a Virtual Organization (VO), which contains two composition modules. The first one interacts with the user to generate a chord progression. The second one, uses social opinion to transform the colors of a digital image into a melody which can be acceptable for the human beings that participate in the society. To evaluate the current approach, a small society of experts qualifies the sound produced following criteria of consonance and musicality.



# Agradecimientos

*Cuando terminas un camino tan arduo como es la escritura de una tesis, uno puede llegar a pensar que se ha llegado tan lejos solo gracias al propio trabajo. Sin embargo, es de necios pensar que he llegado hasta aquí solo por méritos propios. Por eso mismo, este es un buen momento para ser justos con el reconocimiento de aquellos que me han apoyado antes, durante y al finalizar esta tesis.*

*Quiero dar mi más sincero agradecimiento a mi director, Juan Manuel, por brindarme la oportunidad, hace ya casi cinco años, de investigar en algo que además me apasiona, y por aceptarme para realizar esta tesis doctoral bajo su dirección. Su apoyo y confianza en mi trabajo, y todas las facilidades que me ha proporcionado ha sido un aporte esencial, no solamente en el desarrollo de esta tesis, sino también en mi formación como investigadora.*

*Gracias también a mi co-director Marcelo, por enseñarme a escribir mi primer artículo, y por sus novcientos ‘Éstá bien, pero...’ antes de dar una publicación por presentable. Le tengo que agradecer mi nuevo espíritu científico y de autocrítica. A todo el laboratorio del grupo “Sound Music Computing” del Instituto de Ingeniería INESC de Porto, gracias por el calor de la bienvenida y por las oportunidades allí brindadas. En especial, le agradezco a Gilberto la posibilidad de colaborar en tu trabajo, y a Matthew, por acogerme en el grupo de investigación como una más. Espero que este haya sido el primero de muchos más encuentros.*

*Gracias a todo el laboratorio BISITE, a mis compañeros de trabajo, en especial gracias a Sara por esos cafés rápidos, por escuchar mis penas y por las revisiones a altas horas de la noche. Gracias también a mis compañeros, Javi, Chamo, Santos, Gabri, Fran, Fer, Alfonso, Alberto, y al resto del equipo por compartir estos momentos de trabajo, y gracias a las chicas de BISITE, por los pequeños momentos de sonrisas y lágrimas para reírnos de nosotras mismas. Muchas gracias también a Angélica, Belén, María y en general, a los compañeros del departamento. Con ellos comencé mi andadura en la informática, y aunque han pasado de ser mentores a ser compañeros de pasillo, siguen enseñándome muchas cosas.*

*Unas gracias también muy sentidas hacia mis familiares, mi tía Rosa y mis primos Víctor y Raquel, que son como mi segunda familia, y a mis abuelos, en especial a mi abuela María, gracias a su sacrificio hoy he podido llegar hasta donde he llegado. Una mención también muy especial a mis padres,*

*Juanjo y Mercedes, que lo han dado todo por sus hijos para que tengamos lo mejor, y a mis hermanos Juan José y Alicia, que siempre me han apoyado en lo bueno y en lo malo, y se han preocupado por mí. Sin todos vosotros no habría podido escribir una palabra de las que hay aquí ahora.*

*Por último, un muy profundo agradecimiento hacia ti, Fernando, por tu generosidad y por estar ahí en todo momento dándome tu cariño y tu apoyo, por ayudarme a levantar en los momentos más difíciles, y sobre todo, por no perder nunca la fe en mí. Esta tesis lleva también un pedacito de ti.*

# Tabla de Contenido

<b>Declaración de Autoría</b>	<b>iii</b>
<b>Resumen</b>	<b>vii</b>
<b>Abstract</b>	<b>ix</b>
<b>Agradecimientos</b>	<b>xi</b>
<b>Contenido</b>	<b>xiii</b>
<b>Lista de Figuras</b>	<b>xvii</b>
<b>Lista de Tablas</b>	<b>xxi</b>
<b>Glosario</b>	<b>xxiii</b>
<b>Símbolos</b>	<b>xxv</b>
<b>1 Introducción</b>	<b>1</b>
1.1 Hipótesis y Objetivos . . . . .	6
1.2 Motivación . . . . .	7
1.3 Metodología . . . . .	9
1.4 Estructura de la Tesis . . . . .	11
<b>2 Estado del Arte</b>	<b>13</b>
2.1 Creatividad Computacional . . . . .	16
2.1.1 Algoritmos de Generación Creativa . . . . .	17
2.2 Algoritmos Bioinspirados . . . . .	19
2.2.1 Inteligencia en Enjambre (Swarm Intelligence) . . . . .	19
2.2.1.1 Optimización de Colonia de Hormigas (ACO) . . . . .	20
2.2.1.2 Optimización de Enjambre de Partículas (PSO) . . . . .	22
2.2.2 Sistemas Inmunológicos Artificiales (AIS) . . . . .	23
2.3 Sistemas Multiagente y Organizaciones Virtuales . . . . .	25
2.3.1 Clasificación de Agentes . . . . .	25
2.3.2 Sistemas Multiagente (MAS) . . . . .	27

2.3.2.1	Agentes y Creatividad . . . . .	28
2.3.3	Organizaciones de Agentes . . . . .	29
2.4	Máquinas Sociales . . . . .	31
2.4.1	Simulación Social . . . . .	32
2.5	Conclusiones . . . . .	34
<b>3</b>	<b>Codificación de Elementos Musicales en el Tonal Interval Space</b>	<b>35</b>
3.1	Introducción . . . . .	37
3.2	Espacios para la Representación Tonal . . . . .	38
3.3	Tonal Interval Space . . . . .	40
3.3.1	Vectores de Croma . . . . .	41
3.3.2	Tonal Interval Vectors (TIV) . . . . .	42
3.4	Propiedades Geométricas del TIS . . . . .	44
3.5	Conclusiones . . . . .	45
<b>4</b>	<b>Hacia un Modelo Generativo para Crear Acordes</b>	<b>47</b>
4.1	Introducción . . . . .	49
4.2	Construyendo nuestra Función Objetivo . . . . .	51
4.2.1	Relación perceptual $\delta$ . . . . .	51
4.2.2	Disonancia $\xi$ . . . . .	52
4.2.3	Relación con la Escala $\lambda$ . . . . .	52
4.2.4	Relación con la Función Armónica $\phi$ . . . . .	53
4.2.5	Función Objetivo $P$ . . . . .	54
4.3	Explorando Acordes con un Sistema Inmunológico Artificial (AIS) . . . . .	55
4.3.1	Sistemas Inmunes Artificiales y Optimización Multimodal . . . . .	56
4.4	Conclusiones . . . . .	59
<b>5</b>	<b>Sociedades Humano-Agente en un Contexto Creativo</b>	<b>61</b>
5.1	Introducción . . . . .	63
5.2	Arquitectura del Generador de Melodías . . . . .	64
5.2.1	Arquitectura Virtual . . . . .	66
5.3	Descripción del Sistema . . . . .	68
5.3.1	Interacción con el Proveedor . . . . .	69
5.3.2	Extracción del Color . . . . .	69
5.3.3	Generación de Sonido . . . . .	69
5.3.4	Agrupación de Notas . . . . .	71
5.3.4.1	Un Valor de Aptitud para Medir la Calidad del Sonido . . . . .	72
5.3.4.2	Integración del Proceso en Enjambre en un MAS . . . . .	74
5.3.5	Generación de la Melodía . . . . .	76
5.3.6	Roles de Síntesis y Reproducción . . . . .	78
5.3.7	Evaluación Experta . . . . .	78
5.4	Conclusiones . . . . .	79
<b>6</b>	<b>Integrando Aprendizaje Social</b>	<b>81</b>
6.1	Introducción . . . . .	83
6.2	Sistema Social . . . . .	84
6.2.1	Organizaciones Virtuales Aplicadas a Nuestro Problema . . . . .	85
6.3	Proceso de Aprendizaje Basado en Casos . . . . .	88

6.3.1	Midiendo la Similitud en los Colores . . . . .	89
6.3.2	Integrando la Inteligencia de Hormigas en el Proceso de Aprendizaje CBR . . . . .	91
6.3.3	Recogiendo la Opinión Social . . . . .	94
6.3.4	Almacenamiento del Caso . . . . .	94
6.4	Conclusiones . . . . .	95
<b>7</b>	<b>Caso de Estudio: Un Framework Compositor</b>	<b>97</b>
7.1	Introducción . . . . .	99
7.2	Funcionamiento del Framework . . . . .	100
7.3	Evaluación del Generador de Armonías . . . . .	105
7.3.1	Test de Escucha . . . . .	106
7.3.2	Resultados y Discusión . . . . .	107
7.4	Resultados del Sistema de Generación Melódica . . . . .	110
7.5	Resultados del Componente de Aprendizaje . . . . .	115
7.6	Evalando el Framework . . . . .	118
<b>8</b>	<b>Conclusiones</b>	<b>121</b>
8.1	Contribución al Estado de la Técnica . . . . .	127
8.2	Trabajo Futuro . . . . .	129
<b>9</b>	<b>Research Overview</b>	<b>131</b>
9.1	Introduction . . . . .	133
9.2	Related Work . . . . .	135
9.2.1	Computational Creativity . . . . .	136
9.2.2	Swarm Intelligence . . . . .	137
9.2.3	Virtual Organizations of Agents . . . . .	139
9.2.4	Social Computing . . . . .	140
9.2.5	Conclusions . . . . .	140
9.3	The Tonal Interval Space . . . . .	141
9.3.1	Chroma Vectors . . . . .	142
9.3.2	Tonal Interval Vectors (TIV) . . . . .	142
9.4	Modeling the Generation of Chord Progressions . . . . .	145
9.4.1	Constructing Our Objective Function . . . . .	146
9.4.1.1	Perceptual Unrelatedness $\delta$ . . . . .	146
9.4.1.2	Dissonance $\xi$ . . . . .	147
9.4.1.3	Key Relatedness $\lambda$ . . . . .	148
9.4.1.4	Harmonic Function Relatedness $\phi$ . . . . .	148
9.4.2	Searching for Chords with an Artificial Immune System . . . . .	150
9.5	Generating the Melody . . . . .	151
9.5.1	Social System with Virtual Organizations . . . . .	152
9.5.2	System Description . . . . .	153
9.5.2.1	Provider Interaction . . . . .	154
9.5.2.2	Color Extraction . . . . .	154
9.5.2.3	Sound Generation . . . . .	154
9.5.2.4	Notes Grouping . . . . .	157
9.5.2.5	An Aptitude Value to Measure Sound Quality . . . . .	158

---

9.5.2.6	Sound Cloud Generation . . . . .	159
9.5.2.7	Synthesizer and Play Roles . . . . .	161
9.6	Integrating the Learning Core . . . . .	161
9.6.1	Social System . . . . .	162
9.6.2	Learning from previous experiences . . . . .	163
9.6.2.1	Measuring Color similarity . . . . .	163
9.6.2.2	Integrating Swarm Intelligence in CBR Learning Process	165
9.6.2.3	Collecting the Social Opinions . . . . .	166
9.6.2.4	Storing the case . . . . .	167
9.7	Evaluation . . . . .	167
9.7.1	Evaluation of the Chord Progressions Generation Module . . . . .	170
9.7.2	Expert Evaluation in Melodic Generation . . . . .	174
9.7.3	Evaluating the Learning Component . . . . .	176
9.7.4	Evaluating the Framework . . . . .	178
9.8	Conclusions . . . . .	180
9.8.1	Future Work . . . . .	184
<b>A</b>	<b>Una Breve Introducción a la Teoría Musical</b>	<b>187</b>
A.1	Introducción . . . . .	189
A.2	Compases y Figuras Rítmicas . . . . .	192
A.3	El concepto de Tono y Semitono y las alteraciones del sonido . . . . .	193
A.4	Escalas y Tonalidad . . . . .	193
A.5	Intervalos y el concepto de consonancia . . . . .	195
A.6	Acordes y Funciones Armónicas . . . . .	198
	<b>Bibliografía</b>	<b>203</b>

# Lista de Figuras

1.1	Representación general del framework. . . . .	5
3.1	Notación musical para las notas de la escala Do mayor con nombres de letras y la codificación correspondiente debajo de cada nota. . . . .	39
3.2	La red armónica o Tonnetz representa las relaciones tonales como una rejilla. Los nodos son clases tonales y las conexiones corresponden a intervalos musicales [Chew, 2007]. . . . .	40
3.3	Visualización del espacio en seis círculos. Cada círculo muestra la representación de un intervalo musical, correspondiente a $k$ en $T(k)$ . El área gris es el área factible de búsqueda donde se pueden encontrar los puntos $T(k)$ . El ángulo $\varphi(k)$ es la rotación desde la posición $n$ a $n + 1$ en el vector de croma. . . . .	43
3.4	Ilustración de las posiciones relativas de Do mayor (CM) y La menor (Am) en el espacio TIV. . . . .	45
4.1	Interpretación vectorial del cálculo de $\lambda$ para la tonalidad y $\phi$ para la función armónica. (a) Cálculo de $\lambda$ utilizando el ángulo $\theta$ entre el acorde candidato $T_i$ y el centro tonal $T_l$ . (b) Cálculo de $\phi$ utilizando el ángulo $\Phi$ entre el acorde candidato $T_i$ y las funciones armónicas I, IV, y V ambos referenciados por $T_l$ . . . . .	53
4.2	Representación de la propuesta presentada en este capítulo para generar una serie de acordes candidatos dada una secuencia anterior. . . . .	56
4.3	La figura representa una función objetivo hipotética. El objetivo es encontrar los mínimos correspondientes a los buenos individuos que deben ser preservados en nuestra población. Los puntos resaltados con un círculo son los individuos cerca de los mínimos, los puntos resaltados con una X representan los mínimos locales y el símbolo del triángulo representa el mínimo global. Todos ellos pueden ser considerados buenos candidatos en nuestra población. . . . .	58
5.1	Esquema del sistema. Las etapas están subrayadas con los recuadros. . . .	65
5.2	Vista funcional siguiendo la metodología GORMAS. Destacamos las misiones que el sistema debe llevar a cabo. . . . .	68
5.3	El proceso de extracción del color mostrado en esta figura consiste en tres fases: fragmentación, extracción del color principal y dotación del color a los agentes sinestésicos. . . . .	70

5.4	Figura a: Representación cromática de los tonos y colores. Cada área delimitada por las líneas de puntos pertenece a una nota musical siguiendo el sistema de Lagresille. Figura b: Asociación entre las propiedades de los colores y el sonido siguiendo la asociación propuesta entre el color y el sonido. . . . .	71
6.1	Flujo de trabajo del sistema. Las etapas se destacan en los cuadros. . . .	85
6.2	Representación GORMAS para la visión funcional. Esta imagen muestra las misiones de todo el sistema. . . . .	87
6.3	Flujo de trabajo del sistema basado en computación social teniendo en cuenta la organización de agentes. . . . .	88
6.4	Esquema con el proceso CBR llevado a cabo por el Rol Planificador. . . .	89
6.5	Ilustración de la construcción del gráfico $G$ que representa los diferentes caminos que pueden tomar las hormigas. . . . .	93
7.1	Roles del sistema en una Organización Virtual . . . . .	101
7.2	Representación GORMAS para la visión funcional. Esta imagen muestra las misiones de todo el sistema. . . . .	102
7.3	Representación general del framework. Las flechas azules indican la interacción de los usuarios con la parte generadora de armonías. Las flechas rojas indican interacciones de usuarios con la parte generadora de melodías. . . . .	103
7.4	Los acordes utilizados como secuencia de entrada para generar el siguiente candidato. . . . .	106
7.5	Gráfico de la evaluación subjetiva (puntuaciones de los oyentes) en función del valor objetivo. Cada punto representa la evaluación del tercer acorde creado por el sistema, considerando la progresión de acordes dada como entrada. El eje horizontal es el valor objetivo y el eje vertical es la media de las puntuaciones que dieron los oyentes. La desviación estándar se muestra para cada punto usando barras verticales. . . . .	108
7.6	Ejemplo musical de un candidato diatónico y uno no diatónico. Ambos son seleccionados por el algoritmo y clasificados como buenos candidatos en la prueba de audición. . . . .	110
7.7	Colección de imágenes aplicadas al sistema para extraer música de nube de sonido. . . . .	113
7.8	Muestra las tasas de similitud entre los colores y el sonido según los oyentes.114	
7.9	Colección de imágenes utilizadas en el experimento. . . . .	115
7.10	Resultados de la evaluación para cada fragmento extraído de una imagen. Cada número corresponde a una imagen de la Figura 7.9 en el mismo orden.116	
7.11	Comparación entre las validaciones del sistema sin el componente de aprendizaje y el actual. La votación para el componente CBR como el mejor resultado se destaca en azul, la votación para el trabajo sin el aprendizaje se resalta en verde, y finalmente el voto indeciso se representa con una barra roja. . . . .	117
9.1	Vector interpretation of the calculation of $\lambda$ for key and $\phi$ for harmonic function. (a) calculation of $\lambda$ using the angle $\Theta$ between the candidate chord $T_i$ and the center of the key $T_l$ . (b) calculation of $\phi$ using the angle $\Phi$ between the candidate chord $T_i$ and the harmonic functions I, IV, and V both referenced to $T_l$ . . . . .	149

9.2	System schema. The stages are highlighted in the squares. . . . .	152
9.3	The color extraction process shown in this figure consists of three main stages. . . . .	155
9.4	Figure a: Chromatic representation of pitches and colors. Each area delimited by the dotted lines belongs to a musical note following Lagresille system. Figure b: Association between properties of colors and sound following Lagresille's association between color and sound. . . . .	156
9.5	Social workflow. The stages of the system are highlighted in the squares. .	162
9.6	Schema with the CBR process carried out by our Planner Role. . . . .	164
9.7	General representation of the framework. . . . .	170
9.8	The chords used as input sequence to generate the next candidate. . . .	171
9.9	Plot of the subjective evaluation as a function of the objective value. Each point represents the evaluation of the third chord created by the system, considering the chord progression given as the input. The horizontal axis is the objective value and the vertical axis is the mean of the scores the listeners gave. The standard deviation is shown for each point by using vertical bars. . . . .	172
9.10	Shows the similarity rates between colors and sound according to the listeners. . . . .	176
9.11	Results of the evaluation for each fragment extracted from an image. Each number corresponds to one image in Figure 7.7 in the same order. . . .	177
9.12	Comparison between the validations of the first approach and the present one. The vote for the present approach as the best result is highlighted in blue, the vote for the past work is highlighted in green, and finally the undecided vote is represented with a red bar. Note that only the present work considers the validation by experts to improve their results. . . . .	178
A.1	Representación de una serie de notas colocadas de forma ascendente. Las notas que se encuentran en la imagen llevan el nombre justo encima siguiendo las sílabas musicales y el alfabeto musical. . . . .	191
A.2	Representación la relación de las diferentes figuras rítmicas que tienen entre sí. . . . .	192
A.3	Representación de las escalas modelo mayor y menor, junto con los grados de la escala. . . . .	195
A.4	Representación los tipos de intervalos que se pueden obtener al considerar notas dentro de una misma octava. . . . .	197
A.5	Representación los acordes que se pueden obtener dentro de una tonalidad, en este caso Do/C Mayor. También se visualizan los grados de cada acorde y su función. . . . .	199



# Lista de Tablas

3.1	Representación vectorial croma de configuraciones de tono. El ejemplo ilustra la escala de Do mayor (C Mayor). . . . .	41
3.2	Interpretación interválica de $k$ para un vector TIV. “Cons.” Quiere decir “Consecutivos”, mientras que “Sup.” quiere decir “Superpuestos”. . . . .	42
3.3	Pesos aplicados a cada componente o intervalo $k$ para los vectores TIV. . . . .	45
7.1	Comparación de las estadísticas resultantes de la organización de las valoraciones subjetivas de la Figura 7.5 en función de las medidas de distancia de otras representaciones de acordes. Véase el texto para la explicación de cada medida de distancia. . . . .	108
7.2	Las estadísticas resultantes de la organización de las calificaciones subjetivas en función de las medidas de distancia de la representación TIS. . . . .	112
7.3	Comparación de nuestro framework, el sistema JANUS y el trabajo propuesto por Miranda & Kirke’ . . . . .	119
9.1	Chroma vector representation of pitch configurations. The example illustrates the C major scale. . . . .	142
9.2	Intervallic interpretation of $k$ for TIV. . . . .	143
9.3	Comparison of the statistics resulting from organizing the subjective ratings of Figure 7.5 as a function of distance measures from other chord representations. See the text for the explanation of each distance measure. . . . .	173
9.4	Statistics resulting from organizing the subjective ratings as a function of distance measures from the TIS representation. . . . .	175
9.5	Comparison between our proposal, JANUS system and Miranda & Kirke’s work . . . . .	180



# Glosario

<b>VO</b>	<b>V</b> irtual <b>O</b> rganization
<b>MAS</b>	<b>M</b> ulti- <b>A</b> gent <b>S</b> ystem
<b>TIS</b>	<b>T</b> onal <b>I</b> nterval <b>S</b> pace
<b>AI</b>	<b>A</b> rtificial <b>I</b> ntelligence
<b>TIV</b>	<b>T</b> onal <b>I</b> nterval <b>V</b> ector
<b>MIDI</b>	<b>M</b> usical <b>I</b> nstrument <b>D</b> igital <b>I</b> nterface
<b>HAS</b>	<b>H</b> uman <b>A</b> gent <b>S</b> ociety
<b>ACO</b>	<b>A</b> nt <b>C</b> olony <b>O</b> ptimization
<b>AIS</b>	<b>A</b> rtificial <b>I</b> mmune <b>S</b> ystem
<b>PSO</b>	<b>P</b> article <b>S</b> warm <b>O</b> ptimization
<b>CBR</b>	<b>C</b> ase <b>B</b> ased <b>R</b> easoning



# Símbolos

$C(n)$	Vector de Croma
$T(k)$	Vector TIV
$W(k)$	Pesos para los componentes $k$ del vector TIV
$\bar{C}(n)$	Suma de todos los valores de los elementos de $C(n)$
$\varphi(k)$	la rotación desde la posición $n$ a $n + 1$ en el vector de cromas
$L_2$	Norma de un vector
$d^2 [T_1(k), T_2(k)]$	Distancia Euclídea entre dos vectores TIV
$\theta$	Ángulo entre dos vectores TIV $T_1(k)$ y $T_2(k)$
$\delta(T_i, T_{i-1})$	Relación perceptual entre dos vectores TIV $T_i(k)$ y $T_{i-1}(k)$
$\xi(T_i)$	Disonancia de un vector TIV $T_i$
$\lambda(T_i, T_l)$	Relación de un vector TIV $T_i$ con la escala $T_l$
$\phi(T_i, T_l, T_f)$	Relación de un vector TIV $T_i$ con la función armónica $T_f$ .
$P(\delta, \xi, \lambda, \phi)$	Función objetivo que mide la calidad de un acorde respecto del contexto
$n_i$	Nota musical representada, o bien nodo de un grafo $G$
$F(n_a, n)$	Función fitness que mide la calidad de dos sonidos individuales
$C(n_i)$	Medida de consonancia de una nota $n_i$
$D(n_a, n_i)$	Medida de similaridad entre dos notas $n_a$ y $n_i$
$L(n_a, n_i)$	Medida de volumen entre dos notas $n_a$ y $n_i$
$\Theta(n_i, n_l)$	Ángulo entre dos notas $n_i, n_l$
$W_c, W_d, W_l$	Pesos que se atribuyen a cada componente en la función de fitness
	$F$
$p_t$	Posición de una partícula en un enjambre
$v_t$	Velocidad de una partícula en un enjambre
$pb_i$	Mejor posición que una partícula ha encontrado en el enjambre

$pb_n$	Mejor posición encontrada por cualquier partícula en el sistema de enjambre
$A(n_{i-1}, n_i)$	Afinidad entre dos notas $n_i$ y $n_{i-1}$
$nSem$	Número de semitonos entre dos notas $n_i$ y $n_{i-1}$
$\kappa$	Compatibilidad entre dos notas $n_i$ y $n_{i-1}$
$W_t, W_{at}, W_c$	Pesos que se atribuyen a cada componente en la función de compatibilidad $\kappa$
$C = \langle P, S, R \rangle$	Representación de un caso de estudio como una 3-tupla: Dominio, Solución, Revisión.
$P$	Dominio del problema y representa el histograma de color de una imagen.
$S$	Conjunto de pares $N = \langle H, D \rangle$ donde $H$ es la nota musical considerada y $D$ su duración rítmica.
$R$	representa las valoraciones subjetivas de los expertos sobre una melodía.
$Pl$	Planes de los casos originales recuperados.
$Pl'$	Planes generados por un agente a partir de casos originales.
$Re$	Recursos consumidos por un agente al cumplir los planes $Pl'$ .
$Ef$	Función de eficiencia del agente CBR
$f$	Función de fitness o coste que debe ser optimizada.
$G$	Construcción gráfica de los caminos a seguir por una hormiga en una colonia
$N$	Nodos de un grafo $G$ , que se corresponden con las notas musicales
$L$	Enlaces entre los nodos $N$
$f(s, t)$	Función de coste de transición entre dos o más nodos.
$u(n_i, n_j)$	diferencia de las puntuaciones globales de composiciones musicales previas
$l_{ij} \in L$	Conector que une los nodos $n_i$ y $n_j$
$\Delta$	Constante que equilibra la importancia entre $fit_t$ y $u$
$pr$	Probabilidad de que una hormiga pase de la nota $n_i$ a la nota $n_j$ en un momento específico $t$
$A, B$	Constantes que equilibran la ecuación de probabilidad $pr$

---

$N_i^k$	Variable que contiene todos los nodos factibles que se pueden alcanzar desde el nodo $n_i$
$\omega$	Factor de evaporación de feromonas
$\tau_{ij}(t)$	Ecuación que formaliza la actualización de los senderos de feromonas
$\Delta\tau_{ij}^k(t)$	Nivel de feromona depositado en la transición $l_{ij}$



*Para Alicia, la voz de mi conciencia, para Juan José, que siempre me hace reír y para Fernando, porque siempre está. . .*



# Capítulo 1

---

## Introducción

---



VNIVERSIDAD  
D SALAMANCA



# Introducción

---

La colaboración es un concepto cada vez más presente en la vida cotidiana actual, esencial para tener éxito en cualquier tipo de objetivo social. Esto se hace patente, por ejemplo, en el aumento de la literatura sobre los equipos empresariales y de gestión para verificar las mejoras cuando un factor de colaboración está presente en un proceso de negocios. Sin embargo, este no es el único campo donde los equipos, grupos, coaliciones o sociedades cooperan para un propósito común. En educación, hablamos de aprendizaje cooperativo en grupos de estudiantes; en los deportes, la formación y gestión de un equipo es esencial para ganar competiciones; y en otros campos como la ciencia, la mayoría de los resultados importantes provienen de la creación y colaboración de grupos de trabajo y de investigación.

El logro de muchas tareas complejas también requiere la cooperación de más de un individuo. Además, siendo los seres humanos criaturas eminentemente sociales, cualquier actividad desarrollada en grupo es capaz de proporcionar más satisfacción. La tecnología debe ser capaz de apoyar estos procesos de creación y gestión de sociedades, grupos o coaliciones que puedan surgir espontáneamente con el objetivo de maximizar el beneficio esperado.

Algunas tecnologías permiten que los seres humanos trabajen directa o indirectamente con entidades computacionales capaces de actuar de forma autónoma e inteligente, llamados agentes. En particular, los seres humanos y los agentes tienen la capacidad de establecer una serie de relaciones e interacciones para colaborar. El concepto de equipos entre agentes y humanos permite además cumplir objetivos individuales o colectivos que se hayan propuesto dentro de una organización o estructura social. La tecnología de los agentes permite vislumbrar el desarrollo de aplicaciones que apoyan dinámicamente la formación y gestión de aquellas sociedades donde los componentes principales pueden ser humanos, con el apoyo y el beneficio del empleo de agentes de software en un entorno de integración completa. Este tipo de sistemas se denomina sociedades humano-agentes.

Los sistemas de personas y agentes que operan a gran escala ofrecen un enorme potencial, ya que pueden abordar situaciones sociales complejas que son la clave de nuestro futuro [Wang, 2007]. Como aplicaciones posibles dentro de este tipo de sistemas podemos mencionar los mercados virtuales en los que interactúan los agentes y los seres humanos, los entornos de simulación, el área de las aplicaciones sanitarias y médicas, la domótica o la composición automática de música y arte [Gomez de Silva Garza and Gero, 2010, Wang et al., 2007]. Este último campo de investigación supone un desafío, ya que debemos plasmar el comportamiento humano creativo para poder generar contenido valioso desde un punto de vista artístico.

Los diferentes avances computacionales en el campo de la Inteligencia Artificial (AI por sus siglas en inglés) que han ocurrido durante los últimos años han atraído la atención de investigadores con todo tipo de orígenes y motivaciones, creando así campos innovadores que unen conceptos aparentemente dispares como la inteligencia artificial y el arte. De estas dos disciplinas nace el área de la Creatividad Artificial o Computacional, que puede definirse como análisis o síntesis computacional de obras de arte, de una manera parcial o totalmente automatizada [Dartnall, 2013]. El área de Creatividad Computacional ha sido recientemente desarrollada significativamente con la entrada de empresas tan importantes como Google, con proyectos como DeepDream [Google Inc., a], una red neuronal que transforma imágenes o, más recientemente, Magenta [Google Inc., b], el equivalente para la generación de música.

Un factor esencial para crear una interesante obra de arte o una composición musical es la motivación. La motivación puede definirse como el conjunto de los elementos que impulsan al artista a generar contenido creativo, y puede consistir en expresar los sentimientos de los autores, comunicar alguna información o describir un elemento en particular. El resultado final puede depender del estilo artístico que el artista quiera imitar, de su cultura, de sus preferencias personales o del estado de ánimo, y cualquiera de estos factores pueden cambiar a lo largo del tiempo. Este contexto dinámico apunta a la aplicación de sociedades humano-agentes como una opción ideal para crear y desarrollar sistemas abiertos y heterogéneos como los que normalmente se encuentran en el proceso de composición. Estas sociedades humano-agente pueden proporcionar resultados muy interesantes cuando se aplican a un problema de colaboración hombre-máquina para generar música.

En este trabajo proponemos un framework que integre el desarrollo de mecanismos, algoritmos, herramientas y modelos para construir un sistema que contenga una organización abierta en los que actúan los agentes tanto virtuales como humanos para generar música. La propuesta pretende utilizar la motivación que el ser humano proporcione al sistema para iniciar un proceso creativo y transformarla en un resultado

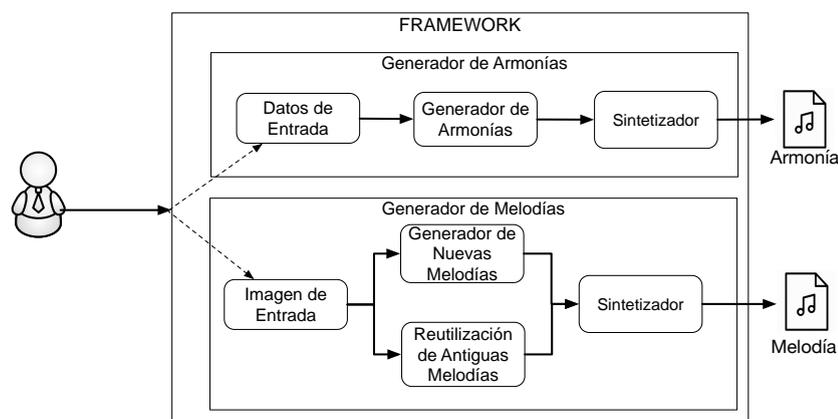


FIGURA 1.1: Representación general del framework.

musical de acuerdo a las preferencias de los distintos usuarios que conforman la sociedad. Es importante destacar que la generación de contenido creativo como la música es un problema desafiante debido al dinamismo del contexto, ya que los usuarios pueden cambiar su motivación en cualquier momento y esperan resultados valiosos según sus criterios, lo cual es también muy subjetivo y puede depender de factores como su cultura o su estado de ánimo.

El framework propuesto se divide en dos partes principales, que pueden ser entendidos como dos sistemas que se pueden integrar posteriormente en dicho framework (que funciona como sistema global) para interactuar con una misma sociedad humana. La primera parte o sistema aborda la generación automática de acordes o armonía, teniendo presente en todo momento la interacción con los usuarios. El framework pretende representar las configuraciones musicales en un espacio geométrico y codificar una función de optimización para seleccionar los mejores candidatos para la generación de los acordes. La segunda parte o sistema del framework consiste en la composición de melodías, utilizando una imagen digital proporcionada por los usuarios como motivación para la creación musical. Esta transformación se realizará utilizando la codificación mencionada en la primera parte, y el resultado también se presentará a los usuarios. Esta parte del framework podrá aprender del contexto social, recogiendo las preferencias de los usuarios para mejorar los resultados musicales finales. Para ello, se desarrolla e integra en nuestra sociedad una arquitectura de Razonamiento Basado en Casos (CBR, por sus siglas en inglés). Inicialmente, el usuario accede a un todo para interactuar y generar, bien armonía o bien melodía. La Figura 1.1 representa el sistema global junto con sus partes principales.

El framework total propuesto hace uso de la computación social para codificar las necesidades humanas, que son interpretadas por la organización virtual de agentes

que comprende la sociedad humano-agente. Además, se integran diferentes estrategias evolutivas para generar una música valiosa para los usuarios.

## 1.1 Hipótesis y Objetivos

Una Sociedad Humano-Agente (HAS, por sus siglas en inglés) puede verse como la siguiente evolución de los Sistemas Multi-Agente (MAS) cuando hay una inmersión de dos niveles de agentes y humanos. En el nivel MAS, el ser humano está situado e integrado en el sistema, y es visualizado como un agente más para el resto de los agentes que participan en el mismo. Mientras, en el nivel humano (desde el punto de vista humano), la interacción se realiza con objetos comunes, ya que los agentes se modelan siguiendo los patrones de diferentes elementos reales. De esta manera, se consigue una integración de ambos tipos de entidades para que el nivel de inmersión del usuario en el HAS sea máximo, minimizando el grado de dificultad de la interacción con el resto de los agentes, así como maximizando el nivel de satisfacción de ese usuario. Una de las metas de este proyecto es profundizar en el desarrollo de las sociedades humano-agente que permiten la inmersión en los dos niveles mencionados anteriormente para generar contenido creativo. Nuestro objetivo es demostrar la posibilidad de crear y proporcionar formación y adaptación a una sociedad de personas y agentes, utilizando como base una organización virtual que permita la entrada y salida de entidades humanas o virtuales, la integración de estas entidades para el logro de objetivos globales, así como una respuesta dinámica a posibles situaciones cambiantes. La capacidad de una sociedad humano-agente para establecer vínculos y lograr objetivos juntos en un entorno dinámico permitirá el desarrollo de sistemas flexibles en los que los individuos son capaces de alcanzar las metas que surgen en la sociedad.

Algunos resultados previos que apoyan la hipótesis de partida están relacionados con el desarrollo de sistemas multiagente para entornos abiertos, en los que los participantes que interactuarán en el sistema final son desconocidos cuando se diseña el sistema [Argente et al., 2011]. Uno de los principales objetivos es el estudio, diseño y desarrollo de nuevos mecanismos, arquitecturas, herramientas y modelos que permitan una organización virtual basada en la tecnología de sistemas multiagente. En esta línea, diferentes grupos de investigación han trabajado en varias áreas para desarrollar modelos en los que agentes virtuales y humanos coexisten e interactúan en un sistema transparente y totalmente integrado. Sin embargo, este tipo de tecnologías nunca se ha aplicado a contextos como la generación de contenido creativo. Por lo tanto, se ha decidido proponer esta tesis basada principalmente en la necesidad de continuar avanzando en estas líneas y aplicarlas a un contexto totalmente novedoso.

Este trabajo de investigación también tiene como objetivo profundizar en el concepto de sociedad humano-agente como un marco basado en la tecnología de sistemas multi-agente para el desarrollo de aplicaciones creativas que generan música. Esto incluye el estudio, definición y desarrollo de métodos, algoritmos y modelos que permiten a las sociedades de este tipo generar composiciones musicales, mediante la colaboración de seres humanos y agentes en un entorno de integración total. Por lo tanto, esta investigación nos permitirá generar una nueva arquitectura o un marco de ejecución para este tipo de sociedad basada en organizaciones virtuales, que facilite ofrecer servicios personalizados y adaptados a la composición musical.

La propuesta incorporará mecanismos que permitan la creación, regulación y adaptación de las sociedades hombre-agente a fin de coordinar sus actividades y acciones para maximizar colectivamente su utilidad en el proceso creativo. Para ello, será necesario crear técnicas de simulación social basadas en agentes [Rodríguez et al., 2011] que permitan el estudio de patrones de comportamiento humano para su validación e integración posterior en los sistemas propuestos. La arquitectura definida tendrá que ser totalmente modular y, por tanto, altamente escalable.

Desde un punto de vista musical, el trabajo tiene como objetivo estudiar y analizar nuevas metodologías y algoritmos para generar música creativa que pueda integrarse en la HAS para facilitar el proceso creativo y la colaboración entre humano y máquina. El trabajo también busca una metodología que traduzca eficientemente el conjunto de elementos que representa la motivación artística en una creación musical. La música compuesta está asociada a una calidad musical que puede ser codificada siguiendo diferentes paradigmas. Esos paradigmas serán objeto de un completo análisis y modificación para adaptarlos al presente problema de composición musical.

Se desarrollarán varios prototipos para la evaluación de la tesis, que serán validados de acuerdo a los objetivos sobre los que se enfocan los prototipos. Cada prototipo ayudará en el logro de los objetivos descritos en esta sección y se integrarán en un sistema completo que permita la interacción de humanos y agentes para crear música.

## 1.2 Motivación

Un sistema multi-agente viene definido por organizaciones que se pueden adaptar dinámicamente de acuerdo con las modificaciones en el entorno o en la especificación de la organización. Esta adaptación dinámica incluye cambios en la estructura y el comportamiento de un MAS, como agregar, eliminar o reemplazar componentes durante la ejecución del sistema sin ser afectados [Dignum, 2003].

Como se señala en Luck et al. [2005], los factores sociales en una organización basada en MAS son cada vez más importantes en un mundo dinámico, siempre conectado y abierto. Esto implica tener el apoyo de agentes (humanos o no) para poder entrar o salir de sociedades virtuales, con una asignación de roles, derechos y obligaciones. De esta manera, el soporte para este tipo de sistemas debe ser considerado por los modelos de reorganización. En este sentido, la adaptación y la evolución del papel que desempeña un agente dentro de una organización no se han considerado de manera general en los enfoques actuales. Los costos asociados a los cambios organizacionales deben tener en cuenta los gastos que dependen de la evolución de las capacidades de los agentes que forman parte del sistema, la evolución de sus relaciones y sus interacciones.

Las plataformas que existían hasta hace pocos años para apoyar las interacciones hombre-agente, se centraron en el diseño de los agentes que participan en el sistema, por lo que no se pueden aplicar directamente para el desarrollo de sistemas abiertos de agentes múltiples donde los participantes del sistema no son conocidos a priori. Uno de los antecedentes más relevantes, la plataforma THOMAS [Argente et al., 2011], está centrada en las interacciones de agentes, independientemente de sus características. Surgió de la necesidad de una plataforma basada en conceptos organizativos capaces de influenciar y limitar el espacio de acciones de los agentes que la unen, sin embargo, no permite la introducción de un sistema dinámico para la generación de contenido creativo.

En la generación de música, es bastante común que exista alguna forma de interacción entre seres humanos y máquinas. Martin et al. [2011] presentó el prototipo de software Toolkit con el objetivo de permitir a los usuarios no técnicos diseñar agentes artificiales e inteligentes para realizar música electrónica en colaboración con un músico humano. Pachet [2003] desarrolló The Continuator, un sistema capaz de interactuar con los usuarios para crear una improvisación de jazz en tiempo real. Thorogood et al. [2012] también presenta un sistema para generar paisajes sonoros basado en tweets sobre noticias recientes. A pesar de estos modelos propuestos de interacción persona-ordenador, las máquinas sociales todavía no se han aplicado en este campo.

Este proyecto nace precisamente del aumento del interés social por la tecnología, de la búsqueda de adaptación de la tecnología al usuario (y no al revés), y del interés emergente por la creatividad artificial, que permite generar música en diferentes contextos de forma automática. El desarrollo de la propuesta es también una oportunidad para adquirir conocimientos en diferentes campos como la creatividad o las tecnologías multi-agente y para aplicar esta diversidad de conocimientos en otros proyectos similares, promoviendo y desarrollando el sector de la innovación.

### 1.3 Metodología

Las actividades que configuran la metodología a utilizar se estructuran de acuerdo con el modelo clásico de fase secuencial que comprende:

1. Análisis y evaluación de propuestas publicadas/desarrolladas.
2. Requisitos del sistema.
3. Diseño de la solución.
4. Realización de la solución/prototipo.
5. Integración de la solución con la desarrollada en otros paquetes de trabajo.
6. Validación de resultados.

En general, las actividades de las tareas se estructuran de la siguiente manera:

- Fase de análisis y especificación de requisitos.
- Diseño y desarrollo del componente o técnicas requeridas para la satisfacción de los objetivos de la tarea.
- Implementación de los componentes y desarrollo de los experimentos con las técnicas propuestas. Las técnicas y diseños propuestos se validarán mediante casos de estudio. Específicamente, el estudio de caso que queremos desarrollar será un sistema creativo basado en agentes capaces de interactuar y adaptarse al comportamiento del ser humano.
- Diseminar los resultados.

Podemos dividir el plan de trabajo en cuatro paquetes principales:

#### **A) Composición de la música a través de mecanismos evolutivos**

El comportamiento de los agentes debe adaptarse dinámicamente a las características de la organización para componer música. Los sistemas de composición musical suelen requerir mecanismos para incorporar reglas, lo que significa que se deben especificar idiomas para la definición de reglas y motores capaces de interpretarlos en tiempo de ejecución. Los algoritmos evolutivos permiten la elaboración de reglas de manera colaborativa entre los agentes de una organización y la definición de un modelo social, facilitando un comportamiento distribuido y adaptativo a las circunstancias de las organizaciones.

Las actividades realizadas en relación con la composición musical son:

- Revisión y estudio de diferentes modelos de composición.
- Definición de los requisitos de nuestro sistema.
- Diseño y evaluación de un sistema que genera contenido creativo mediante la aplicación de algoritmos evolutivos que se integrarán en el comportamiento de los agentes de la sociedad que conforman el sistema.

## **B) Técnicas para el aprendizaje social**

La definición de una sociedad hombre-agente requiere mecanismos que faciliten el aprendizaje social de los comportamientos humanos para reproducirlos en las sociedades de agentes. Para ello se llevaron a cabo las siguientes actividades:

- Revisión del estado del arte de las técnicas que permiten un enlace entre los estímulos y las respuestas hechas por los seres humanos.
- Definición de la ontología que facilita el almacenamiento de los procesos de aprendizaje que intervienen en los procesos de percepción.
- Diseño y evaluación de modelos basados en inteligencia de enjambre que incorporan diferentes modelos de inteligencia artificial para el procesamiento de información que facilita la recreación del comportamiento humano. En particular, el sistema creativo propuesto debe interactuar con los seres humanos para crear “arte”.

## **C) Descripción e implementación de la simulación**

Es necesario realizar un estudio detallado de los diferentes modelos de comportamiento humano para poder llevarlos a cabo en tareas posteriores (simulaciones). A partir de este estudio, los sistemas se diseñarán para modelar y especificar los detalles de todos los componentes y módulos que conforman la simulación del sistema propuesto. Finalmente, se propone el uso de técnicas de extracción de información y patrones de comportamiento basados en los datos obtenidos. Se pretende obtener una serie de técnicas de extracción de conocimiento que permitan aportar conocimientos adicionales a los usuarios y al sistema para la toma de decisiones. El uso de este tipo de técnicas se justifica en el proyecto debido a la cantidad de datos que serán tratados en el contexto.

- Diseño de modelos de comportamiento humano para pasar al modelo computacional.

- Identificación y diseño de los componentes necesarios para el modelo de simulación
- Evaluación del modelo de simulación.
- Modelado de un sistema automático de extracción de información y mecanismos de análisis.

#### **D) Diseminación**

La difusión se centra en la publicación de resultados de investigación en revistas y libros especializados, la presentación de comunicaciones en congresos y talleres especializados.

### **1.4 Estructura de la Tesis**

Para probar la hipótesis de partida y alcanzar los objetivos establecidos, se ha estructurado esta memoria en nueve capítulos.

En el Capítulo 2, se lleva a cabo un estudio sobre el estado del arte de trabajos relacionados con la composición automática de música que de alguna forma hagan uso de colaboración entre hombre y máquina. En ese mismo capítulo se describirán los sistemas multiagente. Se hará un especial hincapié en lo referente al aprendizaje social mediante el estudio de diferentes modelos aprendizaje como son los CBR y la computación social. Además, se dedicará un breve apartado a los principales algoritmos evolutivos empleados en esta tesis.

En el Capítulo 3 se describe un espacio donde se codifican distintas propiedades tonales. En este espacio geométrico, hemos extraído diferentes medidas matemáticas que capturan propiedades musicales de cualquier configuración tonal, tal como escalas, acordes o notas individuales.

En el Capítulo 4 se presenta el sistema generador de armonías, donde se aplican las medidas desarrolladas en el Capítulo 3 a un primer sistema evolutivo para la generación de acordes.

Utilizando también las propiedades desarrolladas en el Capítulo 3, se ha generado un primer prototipo de sociedad humano-agente que permite la colaboración entre humanos y máquinas para generar una composición musical de acuerdo a los cánones tonales. El sistema se detalla en el Capítulo 5 y supone la base para la integración de mecanismos de aprendizaje social, tales como el CBR, todo ello descrito en el Capítulo 6.

Recogiendo todos los resultados principales, el Capítulo 7 describirá la integración de los componentes en el framework, así como los diferentes casos de estudio para validarlo. Por su parte, el Capítulo 8 mostrará las conclusiones alcanzadas a lo largo de la investigación.

Para finalizar, se ha incluido un capítulo en inglés (Capítulo 9) que consiste en un breve resumen del trabajo realizado. Este capítulo se añade con motivo del cumplimiento de la reglamentación para el doctorado europeo.

# Capítulo 2

---

## Estado del Arte

---



**VNiVERSIDAD  
D SALAMANCA**



# Estado del Arte

---

La música se considera un área de investigación interesante en una variedad de campos de investigación porque se ocupa de una actividad humana que es intelectual y emocional. Es un lenguaje universal muy diferente del lenguaje hablado.

Algunos investigadores en el campo de las telecomunicaciones están interesados en el software de música, ya que la música se considera una forma de información y necesitan conocer las características generales de esa información y cómo puede manipularse para conseguir una mayor difusión. Los filósofos están interesados en el software musical porque la capacidad de especificar composiciones musicales (es decir, “componer”) a un nivel más alto que nota por nota, les acercaría un paso más para alcanzar una expresión directa de las ideas musicales. Los ingenieros de software también encuentran retos formidables en áreas como la composición musical, ya que la simulación de esta actividad compleja requiere experiencia en diseño de algoritmos, sistemas expertos, optimización y otras disciplinas relacionadas.

Existen múltiples técnicas para la composición automática de música. Inicialmente, algunos trabajos intentaron generar música basándose en gramáticas de Chomsky [Cohen, 1995, Ebcioğlu, 1990]. Estas gramáticas podían modelar estructuras musicales a pequeña y gran escala. Derivado de las gramáticas, se comenzó a estudiar la aplicación de los modelos de Markov, unos modelos estadísticos que, utilizando probabilidades, también son capaces de generar una composición musical con un entrenamiento previo [Bown, 2011, Colton et al., 2012, Marsden, 2000].

Algunos investigadores utilizaron un enfoque totalmente diferente para componer música, empleando algoritmos bio-inspirados y de aprendizaje automático a partir de datos previos o de relaciones matemáticas de las componentes frecuenciales de la música. Estos métodos abarcan algoritmos genéticos, inmunológicos o redes neuronales [Hoover et al., 2011, Navarro et al., 2015]. Todos estos procesos pueden tener ciertas limitaciones si se tratan de forma aislada. No obstante, existen arquitecturas que permiten la integración de elementos dispares, de forma que se puedan aprovechar las ventajas según

el contexto. Una de estas herramientas son las organizaciones virtuales (VO, por sus siglas en inglés), modeladas como sistemas multiagente (MAS, por sus siglas en inglés), ya aplicadas con éxito en diversos trabajos relacionados con la creatividad Kirke and Miranda [2015], Macedo and Cardoso [2004], Machado et al. [2004].

Para validar la eficiencia de todo sistema, es necesario realizar una evaluación que garantice su calidad. En este caso, es necesaria la participación de ciertas entidades que puedan calificar el resultado obtenido de forma fiable. Debido a la subjetividad del producto generado (música), esto solo sería posible con la participación de una sociedad de expertos musicales, lo que, junto con el sistema software presentado, conformaría una máquina basada en computación social.

En este capítulo, vamos a realizar una revisión acerca del desarrollo creativo y las herramientas que se han utilizado. La primera parte introducirá, por tanto, un breve antecedente sobre la creatividad computacional. La segunda sección presenta información general sobre los algoritmos evolutivos que han sido empleados en la presente tesis para la composición automática de música. El tercer apartado explica conceptos sobre agentes, MAS y VO, mientras que la última sección incluye una breve descripción sobre máquina social utilizando agentes.

## 2.1 Creatividad Computacional

El interés en la creatividad computacional está en aumento en la comunidad científica. Aunque este interés es reciente, existen muchos algoritmos, esquemas y procedimientos para desarrollar una máquina tan inteligente, capaz de crear nuevas ideas o nuevas composiciones artísticas.

La creatividad se considera un componente esencial de la inteligencia humana. En consecuencia, al intentar responder a la pregunta de si las computadoras pueden pensar, es natural preguntarse si las computadoras pueden pensar creativamente.

Algunos investigadores de la AI han tratado de simular la creatividad con las computadoras. Entre los programas más impresionantes desarrollados se encuentran AARON [Cohen, 1995], un programa de pintura que produce obras abstractas y realistas utilizando una pequeña tortuga robótica, combinada con varias estrategias de dibujo, o el sistema Ebcioğlu CHORAL [Ebcioğlu, 1988], que es capaz de producir armonizaciones corales siguiendo el estilo de las corales de J. S. Bach. Después de estos experimentos, y parcialmente debido a la falta de consenso conceptual y teórico, ha habido científicos interesados en exponer nuevas teorías sobre creatividad y conceptos relacionados con ella. Una de los más importantes es Margaret Boden, quien propuso un marco de la

definición de creatividad [Boden, 1987] que sigue teniendo un impacto filosófico en la actual Creatividad Computacional [Colton et al., 2012].

Boden define la creatividad como una habilidad para concebir ideas y artefactos nuevos, sorprendentes y valiosos. Estos tres elementos deben existir en todos los fenómenos creativos. Otro teórico notable, Csikszentmihalyi [1997], defiende el argumento de que la creatividad consta de tres partes principales: el dominio, un conjunto de reglas y procedimientos simbólicos; el campo, que incluye a todos los individuos que actúan en el dominio; y el individuo.

Con respecto al proceso mental que las máquinas pueden simular para ser creativo, Boden está interesado en ejecutar pruebas para determinar si las computadoras pueden concebir ideas que se consideran o parecen ser creativas. En “IA y el hombre natural”, Boden considera que IA es la ciencia del pensamiento y la acción, lo que sugiere que las computadoras son herramientas que se proveen con una inteligencia “similar a la humana”. En este sentido, Lösch et al. [2009] identifican aspectos sobre la creatividad en los individuos, como motivación, externalización, inspiración, etc., para definir requerimientos y funcionalidades de un modelo capaz de potenciar las habilidades creativas del usuario en tareas de diseño.

### 2.1.1 Algoritmos de Generación Creativa

Aunque la creatividad computacional es un campo en reciente expansión, ya existen muchos algoritmos, esquemas y procedimientos para desarrollar una máquina inteligente capaz de crear nuevas ideas o nuevas composiciones artísticas.

Inicialmente, los sistemas basados en gramáticas se utilizaron ampliamente en tareas de composición. Al pensar que la música sigue reglas gramaticales, muchos compositores en el campo de la computación modelaron las relaciones musicales como estructuras gramaticales, representando estructuras musicales [Roads and Wieneke, 1979]. De hecho, Holtzman [1981] creó una gramática musical que genera solos de arpa basados en las limitaciones físicas impuestas a los intérpretes de arpa. Cope [1987] deriva la gramática de los principios lingüísticos para generar música en un estilo particular. Aunque la gramática puede producir un sonido natural, las tareas correspondientes a decidir los aspectos de una estructura musical que debería estar representada, a menudo son complejas y difíciles de modelar [Marsden, 2000]. Hoy en día, hay muchos otros algoritmos que intentan componer música, algunos de los cuales son llamados algoritmos en vivo [Bown, 2011], que derivan de las gramáticas superando algunas de sus limitaciones para generar composiciones musicales.

Uno de los algoritmos más exitosos son los modelos de Markov [Eigenfeldt and Pasquier, 2013]. Los modelos de Markov son modelos estadísticos capaces de predecir notas probables a partir de una melodía ya preestablecida. En este sentido, un interesante estudio a destacar es el de F. Pachet [Pachet, 2003]. Pachet propone un sistema denominado “The Continuator”, que es capaz de construir representaciones operativas de estilos musicales en un contexto de tiempo real. La propuesta se basa en los modelos de Markov para tratar con el ritmo, los lanzamientos y las imprecisiones. El sistema resultante es capaz de aprender y generar música en cualquier estilo y también hace posible crear nuevos modos de reproducción musical colaborativa [Pachet, 2002]. Otros trabajos basados en Modelos de Markov, también usan letras como una variable en sus composiciones, como por ejemplo Monteith et al. [2012].

Con respecto a los algoritmos de bio-inspiración y de aprendizaje automático, Hoover et al. [2011] se centró en la evolución de un solo acompañamiento monofónico para un archivo MIDI. Estos acompañamientos se generan a través de dos funciones, una para el tono y otra para el ritmo, que se representan como una red de producción de patrones de composición (CPPN), un tipo especial de red neuronal artificial (ANN). Las CPPN pueden evolucionar para asumir una topología arbitraria en la que cada neurona tiene asignada una de varias funciones de activación. Mocholi et al. [2012] abordó el problema de la generación de listas de reproducción de música mediante el uso de una colonia de hormigas multicriterio, y López-Ortega and López-Popa [2012] presenta una suite para ayudar a la creación de piezas musicales basadas en fractales, lógica borrosa y sistemas expertos. Otro buen ejemplo es el sistema de reconocimiento para la música occidental propuesto por Mostafa and Billor [2009], basado en algoritmos de aprendizaje de máquina.

Recientemente, Velardo and Vallati [2014] propone un modelo memético para la composición musical, que considera tanto los niveles psicológicos como los sociales. Permite a la gente experimentar la generación de música, pero no se centra en la generación de música clásica. La Sala de Música [Morreale et al., 2014] es un sistema interactivo donde una pareja compone música moviéndose en el espacio. Este es un trabajo interesante, aunque no se centra en la música clásica. AutoRhythmGuitar [McVicar et al., 2014] permite componer ritmos usando algunos parámetros de entrada como acordes y melodía. Dubnov and Assayag [2013] también crea una composición asistida por computadora usando archivos MIDI como entradas para crear una nueva melodía. En este último trabajo. López-Ortega and López-Popa [2012] presenta un sistema experto para ayudar a componer piezas musicales. Este trabajo sigue las reglas del fractal, capaz de componer la armonía atonal. Navarro et al. [2015] presenta un sistema capaz de crear progresiones de acordes basadas en un AIS. Se basa en un

algoritmo evolutivo, que no proporciona la flexibilidad de los sistemas multiagente para adaptar el sistema a diferentes contextos.

## 2.2 Algoritmos Bioinspirados

Los algoritmos bioinspirados son procedimientos capaces de imitar el comportamiento de algún agente que se encuentra en la naturaleza, como insectos o microorganismos, o proceso natural, como la reproducción celular o la sinapsis mental. Este tipo de técnicas algorítmicas han sido avaladas por un amplio número de estudios abarcando campos de investigación muy diversos, como pueden ser la genética, la Medicina, la Robótica, la Física, la Computación o el Diseño Industrial.

Debido a la multitud de técnicas que se agrupan bajo el nombre de algoritmos bioinspirados, se hace inabarcable realizar un estudio exhaustivo de todas ellas en este trabajo. Por ello, las secciones posteriores solo se centrarán en aquellas técnicas relevantes para el objetivo de la tesis, a saber, la inteligencia en enjambre y los algoritmos inmunológicos artificiales (AIS, por sus siglas en inglés).

### 2.2.1 Inteligencia en Enjambre (Swarm Intelligence)

La inteligencia en enjambre estudia el comportamiento colectivo de sistemas descentralizados y auto-organizados. Este concepto se inspira en los sistemas biológicos y se aplica en la inteligencia artificial y en el contexto de los sistemas robóticos y software [Bonabeau et al., 1999].

Estas técnicas se componen a menudo de una población de agentes o partículas que interactúan entre sí y con el contexto, lo que proporciona flexibilidad y robustez [Camazine, 2006]. Los agentes siguen reglas simples, por lo que la interacción local entre ellos conduce a la aparición de un comportamiento global complejo. Los ejemplos naturales son numerosos y diversos entre los insectos o las aves sociales [Camazine, 2006, Detrain and Deneubourg, 2006].

La población del enjambre trata de optimizar una función de acondicionamiento físico mediante la adaptación al medio ambiente. Esta población puede computar datos relacionados con el espacio y el tiempo y responder a factores de calidad del contexto. Su interacción debe desarrollarse en canales amplios y su comportamiento puede ser cambiante según la situación, pero estable a lo largo del tiempo. Estos principios fueron recogidos por [Millonas, 1992].

Como una disciplina científica que incluye campos de investigación tales como la optimización del enjambre o el control distribuido en la robótica colectiva, la inteligencia del enjambre nació de una visión biológicas sobre las habilidades de los insectos sociales para resolver sus problemas de la vida cotidiana [Bonabeau et al., 1999]. Sus colonias, que van desde unos cuantos animales hasta millones de individuos, muestran comportamientos fascinantes que combinan la eficiencia con flexibilidad y robustez [Camazine, 2006]. Desde la gestión del tráfico en una red de forestación, los ejemplos de comportamientos complejos y sofisticados son numerosos y diversos entre los insectos sociales [Detrain and Deneubourg, 2006]. En este trabajo proporcionaremos una breve descripción de los principales y más recientes estudios realizados en el estado de la técnica, optimización de colonia de hormigas (ACO, por sus siglas en inglés) y optimización de enjambre de partículas (PSO, por sus siglas en inglés).

### 2.2.1.1 Optimización de Colonia de Hormigas (ACO)

Este algoritmo [Jaiswal and Aggarwal, 2011, Jr et al., 2015] está inspirado en las acciones de una colonia de hormigas. El objetivo es colocar soluciones óptimas moviéndose dentro del espacio de posibles soluciones. Las hormigas biológicas usan feromonas para registrar sus rutas en el medio ambiente y la calidad de las rutas encontradas. De la misma manera, una simulación de hormigas a través de agentes permite recordar la posición de cada agente y la calidad de las soluciones, posibilitando a las hormigas encontrar una solución óptima en próximas iteraciones.

El ACO por tanto, se basa en el comportamiento de las hormigas. Las hormigas se mueven a lo largo del espacio de búsqueda, que suele ser un gráfico  $G$  que contiene rutas posibles para navegar en el espacio. A través de estos movimientos, las hormigas construyen una solución al problema (un camino posible). Cada solución puede ser evaluada por las hormigas. Si la consideran válida, deben depositar un rastro de feromonas a lo largo de la trayectoria recorrida para así recordar esta nueva solución. Este depósito de feromonas depende directamente de la calidad de la solución encontrada. Esta feromona se evapora durante el tiempo para evitar que el algoritmo converja demasiado rápido a mínimos locales. La feromona puede ser reforzada por las hormigas que evalúan la solución, o por una unidad central que decide si este procedimiento es necesario.

Las hormigas artificiales obtienen soluciones óptimas creando caminos diferentes en el espacio y representando todas las posibles soluciones  $S$ . Ellas registran sus posiciones y la calidad de sus soluciones de acuerdo con una función de fitness, también llamada función de coste o aptitud  $f$  que debe ser optimizada. Esta función puede tener diferentes

restricciones  $c$ . Los valores de calidad influyen en la probabilidad de que un camino sea elegido como la mejor solución [Dorigo et al., 2006]. Por lo tanto, podemos formalizar nuestro problema como un problema de minimización  $(S, f, c)$ . El objetivo es encontrar una solución óptima  $s_i \in S$  según nuestra función de costo  $f$  y las restricciones  $c$ .

El problema debe ser representado como un conjunto de estados posibles  $N = \{n_1, n_2, \dots, n_k\}$ . Cada secuencia finita de elementos  $n_k$  se puede denotar como  $x = \{n_i, n_j, \dots\} \in X$ , donde  $X$  significa todas las secuencias posibles. Dentro de  $X$ , hay algunas secuencias que no cumplen con las restricciones  $c$ .

El conjunto de secuencias posibles  $X'$  pertenece al conjunto de soluciones factibles  $S'$  para que  $X' \subseteq S'$ . La función de coste o de aptitud  $f$  depende de la solución concreta  $s$  y del tiempo  $t$ ,  $f(s, t)$ . La función  $f$  se puede asociar a los distintos estados en las soluciones. Por lo tanto, y teniendo en cuenta la representación, las hormigas artificiales construyen soluciones teniendo en cuenta la construcción gráfica  $G = (N, L)$ , donde  $N$  es el conjunto de estados posibles y los nodos en el gráfico y  $L$  son los enlaces entre los nodos  $N$ .

Nuestros nodos  $N$  y conexiones  $L$  están asociados con un rastro de feromonas  $\tau$  y un valor heurístico  $\sigma$ .  $\tau$  codifica la memoria sobre la calidad de las soluciones encontradas por las hormigas.  $\sigma$  puede ser información adicional sobre el problema. En nuestro caso,  $\sigma$  representa el coste de extender el estado actual [Dorigo and Stützle, 2003].

El comportamiento de las hormigas se puede resumir como sigue. Las hormigas buscan soluciones factibles que minimicen la función  $f$ . Cada hormiga tiene una memoria que almacena los caminos seguidos y sus valores para la función  $f$  para depositar feromonas si es necesario. Las hormigas reciben un estado inicial y tienen que explorar el espacio hasta que se cumple una condición de terminación. En este caso, la condición de finalización es llegar a un último nodo o nota musical, es decir, un estado final indicado en el gráfico  $G$ . Dado un estado  $n_i$ , la probabilidad de seleccionar una ruta que conduce al nodo  $n_j$  depende de la memoria individual, la cantidad de rastro de feromonas en la ruta, el parámetro  $\sigma$  y las restricciones de problemas. Cuando las hormigas han construido una posible solución, pueden volver a recorrer el camino para actualizar los senderos de feromonas si es necesario.

El ACO se ha aplicado con éxito en diferentes contextos para optimizar rutas, a saber, el Problema del Viajante [Bianchi et al., 2002], encaminamiento de vehículos [Bell and McMullen, 2004], optimización de tráfico [Jabbarpour et al., 2014], para Asignación de Tareas (GAP) [Zaza and Richards, 2014] o para reportar puntos de referencia en un protocolo de enrutamiento WSN [Saleem et al., 2015].

### 2.2.1.2 Optimización de Enjambre de Partículas (PSO)

En este algoritmo [Duan et al., 2015, Kennedy, 2010], la mejor solución se representa como un punto en un espacio  $n$ -dimensional. Los individuos comienzan con una velocidad inicial y utilizan un canal de comunicación para intercambiar información relevante. Por lo tanto, las partículas se mueven a través del espacio de la solución y se evalúan periódicamente de acuerdo con una función de aptitud. Sobre las iteraciones, las partículas se acercan a aquellas con los mejores valores de aptitud. Cada individuo se mueve libremente a través del espacio teniendo en cuenta tres factores: el componente de peso de inercia, el componente cognitivo y el componente social [Kennedy, 2010]. En primer lugar, (i) la fuerza de inercia está relacionada con la inercia física y depende de la fuerza anterior aplicada a la partícula; (ii) los componentes cognitivos se denominan fuerzas de atracción entre partículas o grupos de partículas; y finalmente, (iii) el componente social está relacionado con el intercambio de información entre partículas. Dentro del algoritmo, las partículas deben contener varias premisas dentro del sistema  $S$ :

- Permanecer cerca de las partículas vecinas. Cada partícula es atraída por otras partículas dependiendo de ciertos parámetros previamente establecidos. Por lo tanto, las fuerzas de atracción son fundamentales en este tipo de modelo.
- Se mueve hacia el centro de gravedad. Esta regla impide que las partículas se alejen demasiado del centro del sistema.
- Evite las colisiones entre partículas. En este caso, se necesitan fuerzas repulsivas si la distancia entre dos partículas es demasiado pequeña.

Uno de los requisitos para aplicar este procedimiento es tener un punto central como punto de optimización. De ese modo, la posición siguiente  $p_t$  de una partícula  $a$  en el enjambre depende de la posición actual  $p_{t-1}$ , la velocidad actual  $v_{t-1}$ , la mejor posición en la hora actual  $pb_i$ , y la mejor posición encontrada por cualquiera de sus vecinos  $pb_n$ , siguiendo la ecuación [Kennedy, 2010]:

$$p_t = f(p_{t-1}, v_{t-1}, pb_i, pb_n) \quad (2.1)$$

Las partículas necesitan interacción con sus vecinos para conocer el  $pb_n$ . En consecuencia, se necesita una actitud de cooperación para lograr el objetivo de todo el sistema.

Podemos encontrar varios estudios sobre algoritmos PSO en la literatura. Ranjbar-Sahraei et al. [2012] desarrollan un sistema de control para mapear posiciones

utilizando un PSO. Reina et al. [2014] analiza diferentes rupturas de simetría emergente mediante la aplicación de un MAS en el que cada agente se comporta como una abeja. Ghasemi et al. [2015] resuelve el problema del despacho de energía reactiva con un enfoque de optimización de enjambre de partículas.

### 2.2.2 Sistemas Inmunológicos Artificiales (AIS)

El sistema inmune (IS) está presente en los organismos de muchas especies, protegiéndolos de agentes externos nocivos. En el caso de los vertebrados, el sistema inmune está compuesto de diversas moléculas, células y órganos distribuidos en el cuerpo, sin embargo, no están controlados por ninguna entidad central. Todos los elementos que se encuentran en el sistema inmunológico se llaman antígenos. Cuando el antígeno se origina dentro del organismo interno, protege al cuerpo y se llama un auto-antígeno o anticuerpo. Los antígenos procedentes de entornos externos pueden provocar diferentes enfermedades y se denominan no-autoantígenos o simplemente antígenos. El sistema inmunológico es responsable de distinguir entre auto-antígenos y no-autoantígenos, mediante un proceso de reconocimiento de patrones y así atacar a los nocivos [De Castro and Timmis, 2002a].

Inspirado en el comportamiento natural del sistema inmunológico, De Castro and Von Zuben [2000] presenta el algoritmo *CLONALG*, un procedimiento de selección clonal que realiza el reconocimiento de patrones. Este algoritmo permite mutar algunos anticuerpos según su afinidad con un antígeno externo. Para ello, genera copias de los anticuerpos según su afinidad con el antígeno. Las copias son mutadas con una tasa  $\delta$  inversamente proporcional a su afinidad con el antígeno, basado en la ecuación (2.2).

$$\delta = \frac{e^{f_i}}{\beta} \quad (2.2)$$

Donde  $\beta$  es una constante obtenida empíricamente para normalizar el efecto de la tasa de aptitud  $f_i$  obtenida por cada célula. Estos nuevos individuos son agregados a la población en general y reevaluados para ser reproducidos y mutados de nuevo. Así, basado en un comportamiento evolutivo, *CLONALG* aprende a reconocer patrones [De Castro and Von Zuben, 2000].

Con el fin de adaptar el comportamiento inmune a los problemas de optimización, una nueva formulación es desarrollada por De Castro and Timmis [2002a], llamada *opt-aiNet*. La información proporcionada se representa a través de los antígenos para ser reconocidos por los anticuerpos. Definimos fitness o aptitud como la afinidad entre el antígeno y el anticuerpo. Por lo tanto, los valores de alta aptitud reflejan alta afinidad. También,

la aptitud se puede comparar con una métrica de la distancia entre el antígeno y el anticuerpo. Pequeñas distancias representan alta afinidad, mientras que las distancias largas significan baja afinidad.

En la primera etapa, los anticuerpos se generan aleatoriamente. Los antígenos se presentan a los anticuerpos para calcular la afinidad entre ellos. La afinidad se mide con una métrica de distancia tal como la distancia euclidiana, por lo que opt-aiNet es capaz de optimizar funciones en  $\mathbb{R}^N$ . Aquellos con alta afinidad son seleccionados y reproducidos en base a su valor de aptitud de acuerdo con el algoritmo *CLONALG*. Con el fin de preservar la diversidad, los anticuerpos cuya afinidad es inferior a un determinado umbral  $t$  se eliminan de la población. El pseudocódigo correspondiente se presenta en Algoritmo (1).

---

**Algorithm 1** Proceso de Optimización Aplicando Sistema Inmunes

---

```

1: procedure OPT-AINET
2:    $N \leftarrow NumMaxIndividuos$ 
3:    $N_c \leftarrow MaxClonesPorCel$ 
4:    $\delta \leftarrow ParametroMutacion$ 
5:    $t_s \leftarrow UmbralSupresion$ 
6:    $A_b \leftarrow GENERARPOBLACION(N)$ 
7:   repeat
8:      $CALCULARAFINIDAD(A_b)$ 
9:      $MUTARPOBLACION(A_b, \delta, N_c)$ 
10:     $fit_m \leftarrow CALCULARMEDIAAFINIDAD(A_b)$ 
11:    if  $fit_m \leq fit_{mant} + error$  then
12:       $A_b \leftarrow ELIMINARINDIVIDUOS(A_b, t_s)$ 
13:       $A_b \leftarrow GENERARPOBLACION(A_b, 0.3N)$ 
14:    end if
15:     $fit_{mant} \leftarrow fit_m$ 
16:  until Criterio de parada se cumpla
17: end procedure

```

---

Las características más importantes de opt-aiNet son:

- Su capacidad para encontrar varios óptimos de la función objetivo en paralelo, preservando la diversidad de las soluciones. Esto significa que opt-aiNet puede encontrar un conjunto de buenos candidatos diferentes entre sí como solución a un problema de optimización.
- Su capacidad de guardar o memorizar para preservar a aquellos individuos que son lo suficientemente buenos para ser reproducidos y mutados en iteraciones consecutivas.

Opt-aiNet se ha aplicado en diferentes contextos con resultados positivos para resolver problemas combinatorios [Chen et al., 2014, Hsu et al., 2014], para detectar intrusiones

en redes de sensores inalámbricos [Shamshirband et al., 2014] o incluso para generar progresiones de acordes [Navarro et al., 2015].

## 2.3 Sistemas Multiagente y Organizaciones Virtuales

El concepto de agente está presente en diversos campos de estudio, como la psicología, la informática, la sociología, la medicina, la economía, etc., con diferentes puntos de vista y comportamientos. En Ciencias de la Computación, el término de agente es ampliamente utilizado en campos de investigación muy diferentes, tales como la inteligencia artificial distribuida o la interacción hombre-computadora [Corchado et al., 2004].

El concepto de agente ha sido definido varias veces por diferentes autores. Wooldridge and Jennings [1995] considera un agente como un sistema informático que se encuentra en algún entorno y es capaz de actuar de forma autónoma en este contexto para lograr sus objetivos de diseño. Russell and Norvig [1995], que describe la noción de un agente, aparece como una herramienta para analizar sistemas, no una caracterización absoluta que divide al mundo en agentes y no agentes.

Para Russel, un agente puede ser visto como algo que percibe el ambiente a través de sensores y actúa en su entorno a través de actuadores. Por lo tanto, el agente debe ser autónomo, reactivo o racional, y debe ser provisto de algún tipo de inteligencia y habilidades sociales.

### 2.3.1 Clasificación de Agentes

Los agentes pueden clasificarse atendiendo a múltiples parámetros. Para los fines de este trabajo, la clasificación siguiendo el modelo de razonamiento interno es de especial interés. La representación del conocimiento del agente y cómo este utiliza ese conocimiento para resolver problemas, permite dividirlos en dos grupos principales [Bajo et al., 2007]:

- Agentes deliberativos: Son agentes que incorporan el paradigma del pensamiento deliberativo como modelo de razonamiento. Estos agentes hacen una representación simbólica del conocimiento y tienen un modelo de razonamiento sobre ese conocimiento que les permite llevar a cabo la planificación o negociaciones y trabajar de forma coordinada con otros agentes.
- Agentes reactivos: carecen de modelos de representación simbólica del conocimiento, con los que no tienen conocimiento del medio ambiente ni de sí

mismos. Son agentes que reaccionan a los estímulos de su entorno y que para responder a cada estímulo sólo se utiliza el conocimiento actual disponible en el momento en que se produce el estímulo.

Este trabajo se centra en los agentes basados en el modelo deliberativo, y más específicamente en aquellos que se basan en la arquitectura BDI (Belief-Desire-Intention), donde cualquier agente que implementa una arquitectura BDI está dotado de los estados mentales de Creencias, Deseos e Intenciones [Bratman, 1987]. El BDI es una de las arquitecturas más estudiadas para implementar un agente de razonamiento, ya que combina elementos interesantes: un modelo filosófico apreciable de razonamiento humano que es fácil de entender y un número considerable de implementaciones como el control de procesos, la toma de decisiones en los negocios, etc. Se ha desarrollado una semántica lógica y elegante que ha sido aceptada por la comunidad científica [Singh et al., 1999]. Los agentes BDI se caracterizan por sus capacidades, como autonomía, reactividad, proactividad, habilidades sociales, razonamiento, aprendizaje y movilidad entre otros. Las capacidades pueden ser modeladas de diferentes maneras y con diferentes metodologías [Corchado and Laza, 2003, Wooldridge and Jennings, 1995]. Una posibilidad es usar sistemas de razonamiento basado en casos (CBR).

El razonamiento de base de casos (CBR) se fundamenta en la idea de que problemas similares se resuelven a partir de acciones similares, y proporciona al agente un mecanismo para aprender de experiencias anteriores y resolver nuevos problemas. Para un agente BDI-CBR, es lógico resolver un problema aplicando una solución lograda en un caso anterior pero con características similares al caso presente. En este sentido, el agente debe almacenar los resultados anteriores si son buenas soluciones, o descartarlos si la calidad de las mismas es baja. La descripción concreta y formal del problema, su solución y el resultado o evaluación se denomina “caso”, que puede representarse por una 3-tupla,  $\langle \textit{Problema}, \textit{Solucion}, \textit{Resultado} \rangle$ .

Se recomienda que un agente CBR implemente un método en su fase inicial, cuando no se hayan almacenado en la memoria soluciones anteriores para que el agente pueda proponer buenos resultados. Además, siempre se requiere una evaluación para validar una solución. Esta validación debe ser llevada a cabo por un experto en el campo para aumentar la confianza en el aprendizaje [Riesbeck and Schank, 2013].

Para implementar un agente BDI-CBR, debe seguirse un ciclo de vida con cuatro etapas [Aamodt and Plaza, 1994], a saber, Recuperar, Reutilizar, Revisar y Retener. En la etapa de recuperación, el agente busca casos con una descripción similar al problema actual que está intentando resolver. En la etapa de reutilización, el agente CBR

adapta las soluciones recuperadas y así obtener una nueva para el problema actual. El resultado propuesto es evaluado por expertos en la etapa de revisión, para decidir si es lo suficientemente bueno para almacenar la nueva experiencia en la etapa de retención.

El CBR se implementa a través del agente BDI-CBR, que genera acciones utilizando casos anteriores para lograr sus objetivos. Existen varios enfoques que utilizan un CBR en contextos dinámicos. SHOMAS [Bajo et al., 2009] es una plataforma de orientación en centros comerciales que utiliza una arquitectura BDI-CBR. Chamoso et al. [2016] propone una arquitectura CBR con agentes para controlar redes de energía eléctrica. Shen et al. [2015] ayuda a la decisión clínica basada en agentes y CBR junto con conocimiento médico. En la presente propuesta, se aplica un CBR para crear contenido musical.

### 2.3.2 Sistemas Multiagente (MAS)

Diferentes arquitecturas de agentes pueden colaborar para resolver un problema complejo. En tales casos, los agentes están asociados creando un sistema multi-agente (MAS). Es necesario matizar esta afirmación, ya que hablar de que dos o más agentes trabajen juntos es algo muy general. De hecho, para que una asociación entre los agentes que trabajan juntos se considere un MAS, se deben reunir una serie de condiciones. Al menos uno de los agentes debe ser autónomo y debe haber al menos una relación entre dos agentes, lo que significa que al menos uno de los agentes tiene información incompleta o limitada para resolver el problema. Los MAS se caracterizan porque no hay un sistema de control global y porque cada agente se centra en su comportamiento individual. Cada agente puede decidir libremente, dinámicamente, qué tareas realizar y a quién asignar estas tareas [Bajo et al., 2007].

El MAS es considerado como una organización de agentes y definido como una entidad compuesta por diferentes miembros realizando diversas tareas o funciones y estructurada de acuerdo con un patrón de comunicación y topología específicos para lograr el objetivo general de la organización, basado en reglas de comportamiento. Los MAS son una tecnología software motivada por preguntas de investigación fundamental sobre la autonomía, la cooperación, la formación de grupos, etc. Actualmente se están aplicando en una gran variedad de dominios, haciendo contribuciones importantes.

Los MAS se puede clasificar como sistemas abiertos o cerrados. La diferencia fundamental entre los dos es que un MAS cerrado se crea con una estructura fija y objetivos, mientras que en un sistema abierto los agentes pueden entrar o salir del sistema dinámicamente y no están necesariamente diseñados para compartir objetivos comunes. En este trabajo, se presta especial atención a los sistemas abiertos. Los

sistemas abiertos existen en entornos operativos dinámicos, en los que se integran nuevos componentes de forma dinámica. Además, los componentes existentes pueden abandonar el sistema en cualquier momento, y las condiciones de operación pueden cambiar de forma impredecible. Los sistemas abiertos se caracterizan por la heterogeneidad de sus participantes, la confianza limitada, objetivos individuales en conflicto y una alta probabilidad de no conformidad con las especificaciones. El diseño y la construcción de los agentes y su comportamiento subsiguiente dependen no sólo de la arquitectura interna de los agentes, sino también de la plataforma donde se encuentran implementados. Por tanto, la función principal de una plataforma de agente es proporcionar un marco de ejecución para los agentes que forman parte de ella.

Las plataformas que existieron hasta hace unos años se centraron en el diseño de los agentes que participan en el sistema, donde los principales factores a tener en cuenta son aspectos o estructura, funcionalidad, dinamismo, normalización y entorno [Boissier and Gâteau, 2007, Hübner et al., 2010]. Algunas de ellas, debido a su rigidez, no pueden aplicarse directamente para el desarrollo de sistemas multi-agente abiertos donde los participantes del sistema no son conocidos a priori. En vista de ello, en los últimos años han aparecido trabajos de investigación o comerciales para proporcionar precisamente ese marco. De esta manera podemos destacar las propuestas de JADE, FIPA-OS, Grasshopper, Jack, ZEUS, AgentScape, MadKit, EIDE, Jason [Bordini et al., 2005], Instituciones Electrónicas [Esteva et al., 2001], RICA-J [Serrano and Ossowski, 2004], Magentix [Alberola et al., 2007] SIMBA [Julian et al., 2002] o SPADE [Gregori et al., 2006]. La plataforma THOMAS [Argente et al., 2011], centrada en las interacciones de los agentes, independientemente de sus características, surgió de la necesidad de una plataforma basada en conceptos organizativos capaces de influenciar y limitar el espacio de acciones de los agentes que la unen. Por su parte, PANGEA [Zato et al., 2013] permite no solo modelar sistemas multiagente, sino organizaciones virtuales con fines sociales, admitiendo cualquier tipo de lenguaje y plataforma.

### **2.3.2.1 Agentes y Creatividad**

Agentes y creatividad son dos disciplinas que han interactuado en varios casos de estudio en la literatura. Martin et al. [2011] presenta una caja de herramientas para diseñar agentes musicales interactivos. Macedo and Cardoso [2001] propone SC-EUNE, un sistema que genera contenido creativo basado en la curiosidad de los agentes. Machado et al. [2004] desarrolla críticas adaptativas para artistas evolucionistas, basadas también en agentes. Un acercamiento de agentes creativos, llamados agentes motivacionales, fue propuesto y utilizado para explorar entornos desconocidos [Macedo and Cardoso, 2004]. Delgado et al. [2009] propuso Inmamusys, que se basa en agentes y sistemas expertos

para generar contenido musical creativo. Lacomme et al. [2010] propone una actuación artística realizada por un MAS para generar música.

La interacción entre computadoras y seres humanos también juega un papel importante en la generación de música [Maher, 2012]. Esta colaboración se puede modelar fácilmente con MAS. Como ejemplo, Eigenfeldt et al. [2015] describe el musebot y el conjunto musebot para explorar metodologías colaborativas basadas en una combinación de agentes y comunidades para crear una composición colectiva. Kirke and Miranda [2015] propone un sistema multi-agente que genera melodías tonales. Los agentes tienen inteligencia melódica y generan los lanzamientos como resultado de la influencia emocional artificial y la comunicación entre agentes, y la estructura jerárquica de la melodía es el resultado de la estructura social emergente del agente.

### 2.3.3 Organizaciones de Agentes

Hoy en día, hay muchos frameworks multiagente que ayudan y facilitan el trabajo con agentes [Galland et al., 2010, Giret et al., 2010, Howden et al., 2001, Hubner et al., 2007]. Aunque estos frameworks basados en sistemas MAS se consideraron un tema importante hace unos años, ha perdido validez en la actualidad, en un momento en que las computadoras personales, los dispositivos, los teléfonos móviles y los dispositivos similares, han crecido de forma exponencial. Los sistemas MAS abiertos permite modelar este tipo de situaciones, donde el número de agentes y su comportamiento es totalmente dinámico e imprevisible. Cuando un sistema abierto puede ser considerado como entidad social debido al número y tipo de entidades que participan en el mismo, nos encontramos ante una Organización Virtual (VO).

De ese modo, una organización virtual de agentes se define como una entidad social compuesta por un número determinado de miembros que realizan diversas tareas o funciones y que están estructuradas de acuerdo con un patrón de comunicación y una topología específicos para lograr el objetivo general de la organización. Los principales factores son aspectos o estructura, funcionalidad, dinamismo, normalización y medio ambiente [Boissier and Gâteau, 2007, Ferber and Gutknecht, 1998, Hübner et al., 2010].

Esto nos lleva a concluir que las organizaciones virtuales de agentes son una opción ideal para crear y desarrollar sistemas abiertos y heterogéneos como los que normalmente se encuentran en el proceso de composición. El uso de organizaciones virtuales de agentes facilita la incorporación de nuevas técnicas de composición al sistema. Así, la Organización Virtual (VO) permite la creación y desarrollo de sistemas heterogéneos como el que se presenta en esta tesis.

El concepto de sistemas multi-agente y organizaciones o sociedades virtuales necesita una metodología o modelo para ser eficientemente diseñado [Dignum, 2009, Ferber et al., 2009]. Esto ha dado lugar a dos tipos diferentes de métodos de acuerdo con el proceso de desarrollo aplicado [Dignum et al., 2002, Ferber, 2003]. Así, si el proceso se guía por la organización del sistema, decimos que la metodología está orientada a la organización (metodología orientada a la organización). Sin embargo, si el proceso se centra en la especificación de las acciones individuales de los agentes, decimos que la metodología está orientada al agente (metodología orientada al agente).

Entre las propuestas de metodología para diseñar VOs [Argente et al., 2011, Carrascosa et al., 2009, Horling and Lesser, 2004], se eligió la propuesta GORMAS [Esparcia et al., 2014] por estar orientada a la organización y por su flexibilidad y dinamismo. Esparcia et al. [2014] presenta un modelo VO que describe aspectos estructurales, funcionales, normativos y ambientales. Un modelo de organización virtual se compone de cuatro dimensiones:

1. Dimensión estructural: Describe los elementos del sistema y sus relaciones.
2. Dimensión funcional: Detalla la funcionalidad específica del sistema, basada en servicios, tareas y objetivos.
3. Dimensión ambiental: Define cada elemento del entorno al que se puede acceder estableciendo enlaces de comunicación.
4. Dimensión normativa: Describe las reglas de la organización y los objetivos de política que deben seguir los agentes.

Para desarrollar VOs, los roles organizacionales, objetivos, estructura y normas sociales deben ser tomados en consideración. Los roles representan diferentes entidades para cumplir con el propósito de las actividades de la organización. Las acciones de cada rol deben ser especificadas siguiendo los objetivos de la organización. Estos objetivos permiten tomar decisiones sobre cómo diseñar la estructura de la organización. Esta estructura contiene principios que rigen el comportamiento y las relaciones de los agentes. Esto también se detalla en las normas sociales, que contienen el comportamiento esperado de los miembros y las sanciones a aplicar en caso de acciones indeseables.

Las VOs se han aplicado con éxito a diferentes problemas. Videau et al. [2011] desarrolla un VO que apoya al personal médico y establece un sistema inteligente para clasificar un volumen considerable de información. Zato et al. [2013] presenta una VO incorporado para detectar dispositivos con restricciones de recursos y proporciona soporte de usuario en casa como un caso de estudio usando diferentes agentes y sensores. Rodríguez et al. [2015] procesa y fusiona información en redes de sensores, basado en una VO.

## 2.4 Máquinas Sociales

La mayoría de las sociedades hoy se han familiarizado completamente con el concepto de redes sociales. Aunque Internet es un medio joven de comunicación, las redes sociales han revolucionado la forma en que pensamos en menos de una década. Por lo tanto, es lógico que Internet incorpore más rasgos sociales.

En este contexto, hemos visto el surgimiento de la computación social, nuevas formas de computación para hacer frente a estos elementos sociales asistidos por Internet. El término de máquina social se refiere al estudio de sistemas computacionales en los que se tienen en cuenta los comportamientos sociales, o más concretamente cualquier tipo de aplicación computacional en la que el software sirve como intermediario o enfoque de una relación social. Es un término de actualidad debido al auge creciente de Internet como medio de comunicación y de interacción. Este término está muy presente en líneas actuales de investigación tanto en ciencias sociales como en el campo de la informática:

- Ciencias sociales: implementación de convenciones y contextos sociales a través de modelos computacionales donde las personas interactúan.
- Informática: relacionada con la investigación de métodos y técnicas que permiten el desarrollo de sistemas computacionales que muestren comportamientos sociales.

Existen múltiples definiciones de este concepto, propuestas en diferentes trabajos. Schuler [1994] describe la computación social como cualquier tipo de aplicación informática en la que el software sirve de intermediario o enfoque para una relación social, mientras que Charron et al. [2006] lo describe como una estructura social en la que la tecnología pone el poder en los individuos y comunidades, no instituciones. Wang et al. [2007] definen la computación social como la facilitación computacional de los estudios sociales y la dinámica social humana, así como el diseño y uso de las tecnologías TIC que consideran el contexto social. Para Erickson and Kellogg [2000] la informática social implica sistemas que apoyan el comportamiento social entre las personas dentro del sistema y luego hacen uso de ese comportamiento para diversos propósitos. Von Ahn [2009] ve la computación social como una especie de interacción humana en la computadora que combina seres humanos y computadoras para resolver problemas a gran escala que ninguno de los dos puede resolver por sí solo. Más específicamente, esta es una tecnología que soporta cualquier tipo de comportamiento social a través de sistemas computacionales (por ejemplo, blogs, correo electrónico, wiki, redes sociales, etc.). Esta línea de investigación tiene su aplicación en áreas como la simulación social o el comercio electrónico. Atzmueller et al. [2014] presentan un nuevo enfoque para aplicar la computación social a entornos ubicuos. Cheng et al. [2015] utiliza

la informática social para detectar modelos de anomalías para descubrir tendencias de investigación. Zhang et al. [2016] aplica la informática social a los grandes datos móviles. Liu et al. [2016] aplica el paradigma social para controlar la congestión en las redes de tolerancia al retardo.

La computación social también es la base de las últimas tendencias en investigación y desarrollo de sistemas software en los que proliferan los trabajos donde se presta especial atención a la interacción social de los elementos involucrados [Wang et al., 2007]. Este es el caso de las líneas de investigación más actuales dentro de los sistemas multiagente relacionados con la negociación automática, la argumentación, la confianza, la reputación, etc. Casos especiales que se pueden destacar son la gestión de las organizaciones virtuales [Rodríguez et al., 2011], o el aprendizaje social [Garcá-Pardo and Carrascosa, 2011](que actualmente utiliza ideas para adaptar sistemas multiagente que podrían tener su precedente en las ideas utilizadas por los sistemas referenciales basados en el filtrado colaborativo como el utilizado por AMAZON), que han pasado de ser un área de estudio en filosofía, sociología y ciencias humanas en general a ser un término clave en diferentes líneas de investigación sobre sistemas multi-agente en diferentes países.

Como podemos ver, la computación social se ha aplicado con éxito en campos muy diferentes. Queremos demostrar que la computación social también se puede aplicar a un paradigma creativo como la generación de música con buenos resultados.

### 2.4.1 Simulación Social

Robertson and Giunchiglia [2013] indica que la potencia de la computadora social reside en la combinación programable de contribuciones de seres humanos y computadoras. Durante los últimos años, la computación social se ha desarrollado para mejorar las relaciones y comportamientos sociales, capturando información contextual y enriqueciendo el modelo social. Estas características proporcionan herramientas que permiten la cooperación entre seres humanos y computadoras, utilizando Internet para apoyar las infraestructuras de esta comunicación. Este nuevo paradigma de la informática social se puede integrar con diversas tecnologías existentes que proporcionan características para regular patrones y normas dentro de las sociedades. En este sentido, las Organizaciones Virtuales (VO) son un candidato adecuado. Un VO es un sistema abierto formado por la agrupación y colaboración de entidades heterogéneas; hay una separación entre la forma y la función que requiere definir cómo se llevará a cabo un comportamiento [Rodríguez et al., 2011]. Los agentes son un recurso adecuado para implementar un VO dinámico. Modelar un VO como Sistema Multiagente

(MAS) permite describir los elementos de una sociedad específica e incorporar normas, comportamientos funcionales o composiciones estructurales. Debido a que un modelo de interacción social implica muchas características intrínsecas a un sistema abierto, un VO basado en MAS es especialmente apropiado para diseñar una arquitectura apropiada, dada su capacidad de incorporar elementos de una sociedad abierta.

Por tanto, la Simulación Social Basada en Agentes (ABSS) es un área que es responsable de obtener simulaciones de ambientes cambiantes y complejos, entornos en los que se podría utilizar el tipo de sistemas propuesto en este proyecto. ABSS se construye a partir de la intersección de tres campos científicos, como la computación basada en agentes, las ciencias sociales y la simulación por ordenador [Davidsson, 2000]. Los sistemas multiagente tienen características como la autonomía, la reactividad, la proactividad, las habilidades sociales, etc. [Wooldridge and Jennings, 1995], lo que los hace muy adecuados para crear modelos organizativos en los que cada agente pueda tener una correspondencia con un individuo, una organización o un actor de la sociedad que se pretende modelar. Además, las interacciones entre los agentes pueden corresponder a las interacciones que existen en el mundo real [Gilbert, 2004]. Una vez establecido el modelo, se inicia la simulación y se observa el comportamiento del sistema. Si los agentes tienen capacidades adaptativas y de aprendizaje adecuadas, entonces es posible adquirir conocimientos y detectar patrones de comportamiento. Existen muchos modelos de simulación social basada en agentes que tratan de analizar diferentes fenómenos sociales [David et al., 2004, Hare and Deadman, 2004]. Schelling [2006] realizó la primera simulación social basada en agentes. En esta simulación cada persona es modelada como un agente, y las interacciones entre agentes representan procesos sociales relevantes. Epstein and Axtell [1996] desarrollaron el primer modelo de agentes en gran escala, Sugarscape, para modelar sociedades completas y observar diferentes procesos sociales como la muerte, la enfermedad, el sexo, la reproducción, la cultura, los conflictos, las guerras, etc. También David et al. [2004] se centra en modelar sociedades y observar procesos especialmente en el ámbito de la reorganización automática de las organizaciones virtuales.

Una de las grandes ventajas de los sistemas de simulación social basados en agentes es su capacidad para ejecutarse en dispositivos móviles y comunicarse a través de redes inalámbricas [Gilbert, 2004, Li et al., 2008]. Las posibilidades ofrecidas por los sistemas inalámbricos en combinación con sistemas multi-agente son ilimitadas [Wooldridge and Jennings, 1995]. Por un lado, dentro de los flujos de trabajo de computación social organizada, los seres humanos aportan sus competencias, conocimientos y habilidades, junto con sus redes de relaciones sociales y su comprensión de las estructuras sociales. Por otro lado, las TIC pueden buscar y entregar información relevante. Estas habilidades facilitan la construcción de entornos de Inteligencia Ambiental, aportando ubicuidad y

transparencia, ayudando a ser más independientes y a construir sistemas distribuidos más eficientes y capaces de afrontar nuevos problemas [Corchado et al., 2008a,b]. Los seres humanos pueden utilizar esta información dentro de su contexto específico para alcanzar sus metas y, finalmente, para mejorar el ambiente general en el que viven.

Estas tecnologías están evolucionando continuamente y se prevé que tendrán un gran impacto en los próximos años. Algunos ejemplos son el VO y la tecnología de agente [Camarinha-Matos et al., 2004], WSN [Rodríguez et al., 2015], fusión de información (IF) [Rodríguez et al., 2014] o sistemas de localización de interiores [De Paz et al., 2014].

## 2.5 Conclusiones

El estado de la técnica demuestra que las artes visuales pueden aplicarse como una motivación para desarrollar una máquina social creativa, ya que puede ser visto como un problema de colaboración social. Las tendencias recientes en las artes visuales, combinadas con el uso de las redes sociales, dan paso a nuevos modelos de creatividad colaborativa humano-agente. La computación social no se ha aplicado a un contexto creativo como la composición musical de forma relevante, lo que nos alentó a experimentar con una nueva forma de generar arte musical basado en un contexto social agregando un componente humano. Entre los algoritmos de enjambre existentes, PSO y ACO han sido seleccionados como los algoritmos más adecuados para generar contenido musical. La novedad del enfoque propuesto radica en combinar una máquina social incrustada en una Organización Virtual con inteligencia de enjambre para generar un sistema de generación de música.

Además, MAS es una tecnología adecuada para modelar el sistema actual en el que la colaboración hombre-máquina es esencial, ya que parece ser ampliamente utilizado en contextos similares con resultados exitosos. El MAS puede proporcionar flexibilidad para operar en un entorno con elementos cambiantes como los cambios melódicos y las preferencias de los usuarios.

# Capítulo 3

---

## Codificación de Elementos Musicales en el Tonal Interval Space

---



VNiVERSIDAD  
DSALAMANCA



# Codificación de Elementos Musicales en el Tonal Interval Space

---

## 3.1 Introducción

La representación utilizada para codificar las notas juega un papel importante en la generación automática de progresiones de acordes. En particular, las representaciones en las que las relaciones matemáticas como la distancia geométrica entre notas son capaces de reflejar algunas propiedades de la teoría de la percepción o de la música pueden ser particularmente útiles. Estos espacios permiten construir ciertas reglas con las que podemos realizar afirmaciones tales como que los acordes que están más próximos, utilizando alguna medida de distancia, generan mejores transiciones que acordes que están más alejados. Algunos autores han propuesto organizar las notas musicales en espacios geométricos de acuerdo a intervalos musicales y otras propiedades musicales. En estos espacios, las propiedades geométricas como la cercanía reflejan propiedades musicales como la relación perceptual. El Tonnetz [Cohn, 1997] es un ejemplo temprano de tal representación geométrica de las propiedades del sistema tonal porque las pequeñas distancias reflejan relaciones armónicas cercanas. Posteriormente, Chew [2007] propuso un modelo de hélice tridimensional para representar las relaciones tonales. Harte et al. [2006] mapeó vectores cromáticos de 12 bloques al interior de un toroide donde las relaciones armónicas cercanas aparecen como pequeñas distancias. Bernardes et al. [2016] extendió este último espacio, para poder representar las relaciones perceptuales entre diferentes configuraciones a través de la Transformada de Fourier Discreta ponderada (DFT) de unos vectores llamados vectores de croma. En el Espacio de Intervalos Tonales o Tonal Interval Space (TIS), las distancias entre las configuraciones musicales de varios niveles capturan la relación perceptual y las magnitudes indican la consonancia. En algunos trabajos anteriores, la representación

utilizada para codificar configuraciones tonales [Navarro et al., 2015] limitó el sistema a la generación de acordes en la escala de Do mayor y la proximidad en el espacio no refleja las propiedades perceptuales de los acordes. En este trabajo, proponemos utilizar el espacio TIS [Bernardes et al., 2016] para representar configuraciones tonales multinivel tales como las notas individuales, intervalos, acordes y escalas, ya que las propiedades geométricas del espacio están relacionadas con propiedades perceptuales y musicales de estas configuraciones.

Este capítulo tiene como objetivo describir exhaustivamente el Tonal Interval Space (TIS), espacio geométrico que permite representar configuraciones tonales como vectores cuyas propiedades matemáticas son capaces de capturar propiedades musicales. En la primera sección se realizará una breve introducción a algunas propuestas para la representación geométrica de acordes. La Sección 3.3 presenta una formalización matemática del espacio TIS y del proceso de obtención de los vectores de representación en dicho espacio a partir de las codificaciones tonales. La Sección 3.4 recoge las propiedades geométricas que sentarán las bases para, en posteriores capítulos, construir medidas que capturen la calidad musical de las diferentes codificaciones tonales. Por último, se incluirá un apartado de conclusiones a modo de resumen del capítulo.

## 3.2 Espacios para la Representación Tonal

La representación de las notas musicales afecta a la forma en que los acordes, las escalas y muchas otras propiedades musicales están relacionadas entre sí en el espacio subyacente. Perceptualmente, el tono musical varía a lo largo de dos dimensiones llamadas croma y altura tonal [Krumhansl, 1990, capítulo 2]. El croma tonal refleja la circularidad perceptual de las clases tonales o notas mientras que la altura tonal organiza naturalmente los sonidos de grave a agudo. Dentro del contexto de la música tonal, la mayoría de los principios de armonía se basan en la noción de intervalo musical, comúnmente confinada a las relaciones entre las clases tonales, mientras que la tonalidad establece una jerarquía entre los sonidos. La Figura 3.1 ilustra una representación simple para dos octavas de la escala diatónica de Do Mayor (C por su codificación en tablatura americana), donde a cada nota se le asigna un valor numérico entero. Usando esta representación para acordes de tres notas (tríadas), la tónica o primer grado de la escala Do mayor (CM) es el acorde Do mayor  $D_{CM} = [1\ 3\ 5]$  y el grado dominante o quinto es el acorde de Sol mayor (GM)  $D_{GM} = [5\ 7\ 9]$ . Estos acordes se pueden ver como vectores  $D$  en un espacio tridimensional donde los números enteros son las coordenadas. Obsérvese que la dimensión del espacio resultante está dictada por el número de notas, de manera que las clases tonales, los intervalos (díadas), los acordes (tríadas) y las escalas

residen en espacios diferentes. Además, ni las normas de los vectores ni las distancias entre los vectores tienen ninguna relación especial con las propiedades armónicas de los acordes que representan. Por ejemplo, cuando los acordes que comparten una estrecha relación armónica en el sistema tonal se encuentran en puntos cercanos en el espacio subyacente, la proximidad puede usarse como criterio para la progresión automática entre acordes. Un algoritmo inteligente para generar progresiones de acordes podría añadir iterativamente el acorde más cercano al previamente agregado.

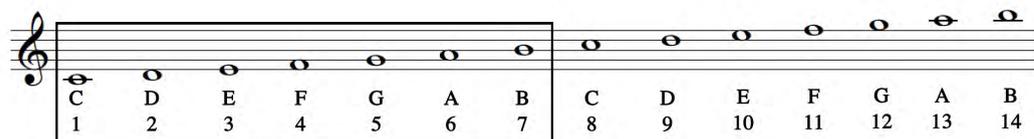


FIGURA 3.1: Notación musical para las notas de la escala Do mayor con nombres de letras y la codificación correspondiente debajo de cada nota.

La Red de armónicos o Tonnetz [Chew, 2007, Cohn, 1997] organiza clases tonales usando los intervalos musicales de quintas perfectas (5P), terceras mayores (3M) y terceras menores (3m), resultando en la cuadrícula mostrada en la Figura 3.2. Los pasos horizontales corresponden a 5P, diagonales a 3M y 3m. La estructura geométrica de la red armónica refleja propiedades de nivel superior del sistema tonal porque las pequeñas distancias reflejan relaciones armónicas cercanas. Por ejemplo, las tríadas son triángulos que apuntan hacia arriba y las clases tonales que pertenecen a una determinada tonalidad comparten una conexión [Chew, 2007]. Por lo tanto, la organización geométrica de las clases tonales en el Tonnetz puede ser explotada en la generación automática de progresiones de acordes. Por ejemplo, los acordes CM y GM están cerca y todos los acordes de Do mayor están dentro de la región ocupada por la tecla Do mayor, de modo que las distancias y las formas se pueden usar como métrica al comparar las notas, los acordes y las escalas. Hay que tener en cuenta, sin embargo, que el Tonnetz es menos adecuado para comparar diferentes niveles de configuración como acordes y escalas.

Chew [2007] propuso envolver el Tonnetz como un tubo con la línea de quintas convirtiéndose en una hélice en la superficie y las terceras mayores alineadas directamente encima una de la otra. La Espiral de Chew permite la visualización y representación de notas, acordes y escalas como puntos en la superficie y dentro de la configuración en espiral. Por ejemplo, las tríadas se representan como un punto en la cara del triángulo esbozado por sus componentes tonales (las notas), de modo que cada tríada es una combinación convexa de su fundamental, quinta y tercera. La Espiral es capaz de representar las notas, acordes y escalas como puntos en el mismo espacio, permitiendo una comparación a través de diferentes niveles. Además, se pueden calcular distancias entre puntos para reflejar las relaciones deseadas entre ellas. Así, la distancia entre el

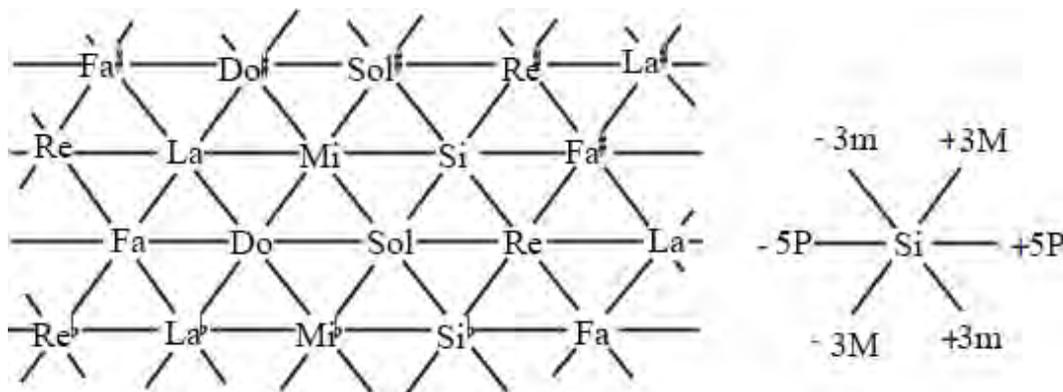


FIGURA 3.2: La red armónica o Tonnetz representa las relaciones tonales como una rejilla. Los nodos son clases tonales y las conexiones corresponden a intervalos musicales [Chew, 2007].

centro de la tríada de Do mayor y sus notas componentes puede reflejar las propiedades del sistema tonal.

En entonación justa, el Tonnetz y consecuentemente la Espiral se extienden infinitamente. Asumiendo la equivalencia enarmónica y de octava, la espiral se curva en un hipertoroide que mapea el croma tonal en su superficie. Harte et al. [2006] se propone representar y visualizar configuraciones tonales dentro de un espacio de seis dimensiones contenido por la superficie de este hipertoroide usando un solo punto de centroide para describir clases de una sola nota o colecciones tales como acordes y escalas. En este espacio, una pequeña distancia entre los centroides tonales refleja estrechas relaciones armónicas entre las configuraciones tonales. Esta representación de seis dimensiones sólo utiliza quintas perfectas, terceras mayores y menores para representar todos los intervalos posibles. Bernardes et al. [2016] proponen vectores de intervalos tonales o Tonal Interval Vectors (TIV) como una representación dodeca-dimensional del croma tonal capaz de manejar todos los intervalos cromáticos. En este trabajo, usamos TIVs para representar configuraciones tonales con un número de hasta doce notas, desde clases tonales basadas en notas individuales hasta acordes y escalas.

### 3.3 Tonal Interval Space

El Espacio de Intervalos Tonales o Tonal Interval Space (TIS) Bernardes et al. [2016] mapea vectores cromáticos de doce dimensiones a Vectores de Intervalos Tonales o Tonal Interval Vectors (TIVs) de valor complejo mediante la Transformada de Fourier Discreta (DFT). El vector de croma puede usarse para representar diferentes niveles de configuraciones tonales tales como notas individuales, intervalos, acordes y escalas. Todas estas definiciones y conceptos se pueden encontrar también en el Apéndice A.

Además, el análisis de Fourier ha sido ampliamente utilizado para explorar las relaciones armónicas entre las clases tonales. Los TIVs Bernardes et al. [2016] son útiles para explorar las propiedades del sistema tonal debido a que las distancias euclidianas entre TIVs capturan la proximidad perceptual entre configuraciones tonales y la magnitud de TIVs refleja la disonancia de la configuración tonal representada.

### 3.3.1 Vectores de Croma

TABLA 3.1: Representación vectorial croma de configuraciones de tono. El ejemplo ilustra la escala de Do mayor (C Mayor).

Posición $n$	Vector de Croma $C(n)$											
	0	1	2	3	4	5	6	7	8	9	10	11
Clase Tonal	C	C $\sharp$ \D $\flat$	D	D $\sharp$ \E $\flat$	E	F	F $\sharp$ \G $\flat$	G	G $\sharp$ \A $\flat$	A	A $\sharp$ \B $\flat$	B
Valor	1	0	1	0	1	1	0	1	0	1	0	1

El vector de croma o simplemente croma  $C(n)$ , representado en la Tabla 3.1, contiene una posición o dimensión para cada una de las doce clases tonales de la escala cromática. En este trabajo utilizamos la notación musical simbólica, por lo que el vector de croma es binario y la presencia de una nota se indica con el valor 1. La Tabla 3.1 supone una equivalencia enarmónica y de octava característica de una afinación de temperamento igual. No hay información sobre la altura codificada en  $C(n)$ . En consecuencia, la octava no puede ser representada por  $C(n)$  con codificación binaria porque todas las octavas se colapsan en una. La inclusión de la octava se estudiará conjuntamente con la mejora de la arquitectura social en el Capítulo 6.

El vector de croma permite la representación de configuraciones tonales multi-nivel indicando la presencia de las respectivas clases tonales (pitch classes en inglés). Por ejemplo,  $C(n)$  para la clase tonal Do, donde Do es considerada como nota individual sería  $C_C(0) = [1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0]$ , para el acorde Sol mayor (G Mayor) es  $C_{GM}(2, 7, 11)$  y para la escala diatónica de la escala de Do mayor se muestra en la Tabla 3.1.

El vector de croma  $C(n)$  ocupa un espacio de doce dimensiones independientemente de la configuración de altura que representa. Sin embargo, las propiedades geométricas del espacio considerado por el vector de croma no captan las propiedades armónicas o musicales de las configuraciones tonales que representa. En otras palabras, los vectores cromáticos  $C(n)$  que representan relaciones armónicas perceptualmente similares no están necesariamente cercanos en el espacio geométrico, calculándolo por ejemplo, con la medida aritmética norma  $L_2$ . Sin embargo, el DFT es capaz de mapear los vectores cromáticos a vectores de intervalo tonal (TIV) en un espacio que exhibe propiedades útiles para explorar las relaciones armónicas del sistema tonal.

### 3.3.2 Tonal Interval Vectors (TIV)

Los TIVs  $T(k)$  se calculan a partir de la DFT del vector de croma  $C(n)$  de acuerdo a la Ecuación 3.1.

$$T(k) = W(k) \sum_{n=0}^{N-1} \bar{C}(n) e^{-\frac{j2\pi}{N}kn}, \quad k \in \mathbb{Z} \quad \text{with} \quad \bar{C}(n) = \frac{C(n)}{\sum_{n=0}^{N-1} C(n)}, \quad (3.1)$$

donde  $N = 12$  es la dimensión del vector cromático y  $W(k)$  son pesos derivados de las puntuaciones de disonancia empírica de las díadas usadas para ajustar la contribución de cada dimensión  $k$  del espacio Bernardes et al. [2016].  $T(k)$  utiliza  $\bar{C}(n)$ , que es la suma de los valores de  $C(n)$  normalizado por el componente DC de la Transformada de Fourier  $T(0) = \sum_{n=0}^{N-1} C(n)$ , para normalizar la representación de todas las configuraciones tonales de niveles múltiples representadas por  $C(n)$ . La Sección 3.3.1 señalaba que  $C(n)$  puede representar cualquier subconjunto de clases tonales, tales como intervalos, acordes y escalas. Equivalentemente,  $T(k)$  permite comparar entre diferentes niveles jerárquicos representando diferentes configuraciones multinivel en el espacio.

El Tonal Interval Space utiliza la interpretación mostrada en la Tabla 3.2 ilustrada en la Figura 3.3, que representa el espacio TIS mediante seis círculos. Cada círculo de la Figura 3.3 corresponde a cada una de las componentes del vector  $T(k)$  cuando  $1 \leq k \leq 6$  de acuerdo a la Ecuación 3.1. Los números enteros alrededor de cada círculo corresponden a las posiciones en el vector croma  $C(n)$ ,  $0 \leq n \leq N - 1$  para  $N = 12$ . Para un  $k$  fijo, la Ecuación 3.1 genera  $N = 12$  puntos igualmente espaciados alrededor del círculo unitario separado por  $\varphi(k) = \frac{-2\pi K}{N}$ . El ángulo  $\varphi(k) = \frac{-2\pi k}{N}$  se ilustra en la Figura 3.3 para  $1 \leq k \leq 6$ .

TABLA 3.2: Interpretación interválica de  $k$  para un vector TIV. “Cons.” Quiere decir “Consecutivos”, mientras que “Sup.” quiere decir “Superpuestos”.

	Vector TIV					
Posición $k$	1	2	3	4	5	6
Pasos $n$	Cons.	Sup.	Sup.	Sup.	Cons.	Sup.
Intervalo	2m \ 7M	TT \ 4A \ 5D	3M \ 6m	3m \ 6M	4P \ 5P	2M \ 7m

La representación interválica en la Tabla 3.2 surge de la interpretación de estos puntos como intervalos musicales usando el vector de croma  $C(n)$  como en la Sección 3.3.1. Para  $k = 1$  y  $k = 5$ , consideramos los intervalos formados entre las posiciones *adyacentes*, por lo que para  $T(1)$  las posiciones adyacentes saltan 1 paso (como 1 y 2). Los intervalos que aquí se representan son segundas menores (2m) o séptimas mayores (7M) si calculamos su complementario. La misma interpretación se aplica para  $T(5)$ , donde las posiciones

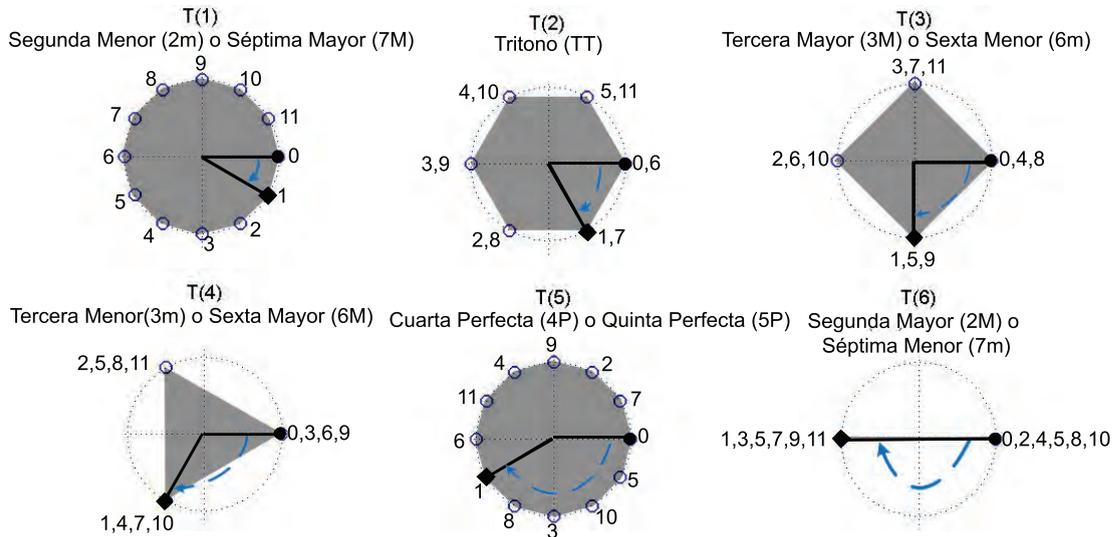


FIGURA 3.3: Visualización del espacio en seis círculos. Cada círculo muestra la representación de un intervalo musical, correspondiente a  $k$  en  $T(k)$ . El área gris es el área factible de búsqueda donde se pueden encontrar los puntos  $T(k)$ . El ángulo  $\varphi(k)$  es la rotación desde la posición  $n$  a  $n + 1$  en el vector de cromas.

adyacentes ahora saltan 5 pasos (tales como 1 y 6) y corresponden a los intervalos de cuarta justa o perfecta (4P) o su complementario, la quinta justa o perfecta (5P). Para  $k = 2$ ,  $k = 3$ ,  $k = 4$ , y  $k = 6$ , consideramos los intervalos formados entre las posiciones de *superposición*. En particular, para  $T(2)$  las posiciones superpuestas saltan 6 pasos (como 1 y 7), y corresponden a los intervalos tritono (TT), a saber, cuarta aumentada (4A) y quinta disminuida (5D). Para  $T(3)$ , las posiciones superpuestas saltan 4 pasos (como 1, 5 y 9), y son musicalmente interpretados como intervalos de tercera mayor (3M) o su complementario (6m), y así sucesivamente.

El Tonal Interval Space presenta propiedades musicalmente interesantes [Bernardes et al., 2016], tales como la capacidad de representar y comparar las configuraciones tonales multi-nivel de una forma matemática relativamente sencilla. Una de las propiedades musicales más importantes de los vectores TIV ilustradas en la Sección 3.3.2 es que las configuraciones tonales que comparten algunas relaciones armónicas tienen relaciones de cercanía en el TIS. Por ejemplo, los acordes de la escala principal de Do Mayor se representarán cerca del centro tonal de la escala. Por otro lado, la magnitud de  $T(k)$ , que es la distancia de cada vector TIV al origen, puede usarse como medida de consonancia. La siguiente sección describirá cómo codificar una función de aptitud basándose en las medidas del TIS que capta la calidad musical de los acordes.

### 3.4 Propiedades Geométricas del TIS

Las propiedades del TIS derivadas de medidas matemáticas, como la distancia euclídea o el ángulo, son la principal herramienta para medir la calidad musical de cualquier configuración tonal.

Esta sección ilustra las propiedades del TIS basadas en el álgebra compleja de la Ecuación (3.1), donde  $T(k) \in \mathbb{C}^M$ ;  $M = 6$ . Esto es,  $T(k)$  es un vector de valor complejo en el espacio de seis dimensiones calculado con la Transformada de Fourier cuando  $1 \leq k \leq 6$ . Usando la norma  $L_2$  en  $\mathbb{C}^M$ , podemos definir el producto escalar entre  $T_1(k)$  y  $T_2(k)$ , y la distancia euclídea entre  $T_1(k)$  y  $T_2(k)$  como sigue

$$T_1(k) \cdot T_2(k) = \|T_1(k)\| \|T_2(k)\| \cos \theta = \sum_{k=1}^M T_1(k) \overline{T_2(k)}, \quad (3.2a)$$

$$d^2 [T_1(k), T_2(k)] = \|T_1(k) - T_2(k)\|^2 = \sum_{k=1}^M |T_1(k) - T_2(k)|^2 \quad (3.2b)$$

donde  $M = 6$  es la dimensión del espacio, la Ecuación (3.2a) es el producto escalar,  $\theta$  es el ángulo entre  $T_1(k)$  y  $T_2(k)$ , y  $\overline{T_2(k)}$  denota el conjugado complejo de  $T_2(k)$ . La Ecuación (3.2b) es la distancia euclídea entre  $T_1(k)$  y  $T_2(k)$ . La Figura 3.4 muestra los vectores TIV para los acordes de La (A) mayor y menor para ilustrar sus posiciones relativas en el espacio TIS.

Las Ecuaciones 3.2a y 3.2b expresan la capacidad del espacio TIS para representar similitudes perceptivas usando relaciones algebraicas. La Ecuación 3.2a permite calcular el ángulo  $\theta$  entre vectores TIV, que puede usarse como medida de similitud, donde ángulos menores indican un mayor grado de similitud. La Ecuación 3.2b puede usarse con el producto escalar para estimar la similitud perceptual entre dos vectores TIV porque vectores perceptualmente similares producen distancias más pequeñas que aquellos vectores disimilares.

La Ecuación 3.3 mide la magnitud (norma) de  $T_1(k)$ , que es la distancia desde el origen del espacio TIS utilizado como medida de consonancia [Bernardes et al., 2016].

$$\|T_1(k)\|^2 = T_1(k) \cdot T_1(k) = \sum_{k=1}^M |T_1(k)|^2, \quad (3.3)$$

Los vectores TIV con menor magnitud (por ejemplo, que están cerca del origen) corresponden a conjuntos de clase tonales que son más disonantes que los de mayor



FIGURA 3.4: Ilustración de las posiciones relativas de Do mayor (CM) y La menor (Am) en el espacio TIV.

magnitud. Los pesos  $W(k)$  en la Ecuación 3.1 ajustan la importancia de los intervalos para cada  $k$  en el vector  $T(k)$ . La Tabla 3.3 muestra los valores de  $W(k)$  ajustados para aproximar la consonancia percibida de clases interválicas de temperamento igual derivadas de valoraciones empíricas de tonos complejos en la región de tonos centrales [Huron, 1994]. En este trabajo, los pesos  $W(k)$  fueron elegidos para enfatizar los intervalos 5P/4P, 3m/6M, 3M/6m y 2M/7m, ya que nos encontramos ante un caso de estudio para generar progresiones de acordes siguiendo el estilo de composición tonal diatónico.

TABLA 3.3: Pesos aplicados a cada componente o intervalo  $k$  para los vectores TIV.

	Peso $W(k)$ para cada intervalo $k$					
Intervalo	2m/7M	TT/4A/5D	3M/6m	3m/6M	4P/5P	2M/7m
$W(k)$	0.0187	0.0416	0.203	0.221	0.238	0.121

### 3.5 Conclusiones

Como hemos visto a lo largo de este capítulo, existen diferentes modelos geométricos para representar acordes, notas o escalas. Entre ellos, se ha escogido el Tonal Interval Space como el espacio ideal para codificar nuestras configuraciones tonales. Los motivos de esta

elección están fundamentados en la sencillez de su uso a través de la transformada de Fourier de los vectores de croma de sonido, en las relaciones existentes entre las medidas geométricas y la calidad sonora de cada una de las configuraciones y en la posibilidad de comparar sonidos de distinta configuración tonal, como pueden ser notas, intervalos, acordes o escalas de forma simultánea y sin ningún tratamiento previo adaptado al tipo de sonido. A lo largo de este trabajo se utilizarán la distancia euclídea y la medida del ángulo entre vectores, definidas y formalizadas en el presente capítulo como medidas base para desarrollar futuras funciones de optimización o de captura de la calidad de una configuración tonal.

# Capítulo 4

---

## Hacia un Modelo Generativo para Crear Acordes

---



VNiVERSiDAD  
D SALAMANCA



# Hacia un Modelo Generativo para Crear Acordes

---

## 4.1 Introducción

En composición musical, la generación de música tal y como hemos visto en el Capítulo 2, es un gran punto de interés para la comunidad. En este capítulo se detallará el diseño e implementación de la parte de nuestro framework encargada de generar una estructura armónica basada en progresiones de acordes.

En este sentido, existen soluciones tecnológicas previas que algunos autores han propuesto para generar automáticamente composiciones armónicas basadas en acordes. Estas herramientas no sólo requieren estrategias para la construcción de acordes, sino también la codificación de un conjunto de reglas que gobiernan las transición entre un acorde y otro. Además, los sistemas generativos para las progresiones automáticas de acordes deben capturar las idiosincrasias técnicas del sistema tonal, al tiempo que ofrecen libertad creativa y expresión personal.

Se han diseñado múltiples propuestas para la generación automática de progresiones de acordes siguiendo diferentes paradigmas, incluyendo el aprendizaje estadístico, las gramáticas cíclicas, [Ebcioğlu, 1990, Steedman, 1984], o los principios biológicos [Fukumoto, 2014, Moroni et al., 2000, Navarro et al., 2015].

Un enfoque conocido para generar progresiones de acordes es el aprendizaje estadístico de secuencias de acordes de corpúsculos musicales para alimentar diferentes algoritmos de inferencia tales como modelos de Markov de orden variable [Eigenfeldt and Pasquier, 2005], algoritmos de árbol de decisión [Paiement et al., 2005], modelos estadísticos y evolutivos [Manaris et al., 2007] y las estrategias de correspondencia de patrones [Cope, 1992].

Un enfoque alternativo al aprendizaje estadístico es derivar reglas de los principios de la teoría de la música para diseñar algoritmos basados en principios determinísticos y probabilísticos. Dentro de este enfoque, podemos destacar los modelos computacionales de Anders and Miranda [2009] y Bernardes et al. [2016] que crean progresiones de acordes siguiendo principios de Schoenberg [Schönberg, 2006] y Riemann [1893]. Estrechamente ligado a este enfoque está el uso de gramáticas (generativas) para organizar los lexemas musicales. CHORAL [Ebcioğlu, 1990] es una notable aplicación de este concepto para la armonización de los corales al estilo de Johann Sebastian Bach. El sistema se basa en las gramáticas y contiene alrededor de 350 reglas que representan el conocimiento musical desde múltiples puntos de vista de los corales de Bach, como la construcción del acorde, las líneas melódicas de las partes individuales y las diferentes voces. Steedman [1984] exploró estrategias similares para generar progresiones de acordes de jazz. Delgado et al. [2009] propuso Inmamusys, un marco para la composición musical basado en sistemas multiagente y un conjunto de reglas. López-Ortega and López-Popa [2012] propusieron una herramienta para ayudar en la construcción de piezas musicales usando un conjunto de reglas basadas en fractales y lógica borrosa.

El uso de principios biológicos o evolutivos aplicando conjuntamente una función objetivo que se debe optimizar ha sido otra estrategia recurrente utilizada en la generación automática de progresiones de acordes, en particular a través del uso de algoritmos genéticos. En una de las primeras obras para aplicar la computación evolutiva a la composición algorítmica, Moroni et al. [2000] propuso Vox Populi, que utiliza algoritmos genéticos para evolucionar una población de acordes maximizando múltiples criterios musicales. Fukumoto [2014] también aplica algoritmos genéticos para crear automáticamente progresiones de acordes según el sentimiento del usuario. Herremans and Sörensen [2013] propuso un sistema experto para componer contrapunto basado en diferentes funciones de optimización. Recientemente, se ha empleado un Sistema Inmunológico Artificial (AIS) para generar iterativamente el siguiente acorde en una secuencia dada optimizando una función objetivo [Navarro et al., 2015].

La principal contribución del capítulo es una función multi-objetivo, que se integrará en el sistema para captar las propiedades del sistema tonal, tales como la falta de relación perceptual, la disonancia, la relación con la tonalidad y la relación de funciones armónicas codificadas como relaciones geométricas en el TIS como los ángulos y las distancias. Utilizamos un sistema inmune artificial (AIS) llamado opt-aiNet [De Castro and Timmis, 2002b] para explorar el TIS y encontrar múltiples mínimos locales de la función objetivo en paralelo mientras se preserva la diversidad. Los mínimos locales de la función objetivo son posibles acordes candidatos en diferentes regiones del TIS, representando múltiples soluciones que son diferentes entre sí por tener características musicales distintas.

Este capítulo está organizado de la siguiente manera. En la Sección 4.2, describimos las propiedades del TIS que se consideran en este trabajo y la función objetivo que se ha diseñado para medir la calidad del acorde. Posteriormente, en la Sección 4.3 se describe el sistema general con la integración de la función objetivo y el algoritmo AIS. Finalmente, la Sección 4.4 presenta las conclusiones.

## 4.2 Construyendo nuestra Función Objetivo

Con el fin de capturar la calidad musical de las configuraciones tonales, se ha diseñado una función multi-objetivo utilizando la definición matemática de las propiedades derivadas del TIS. Esta función se divide en cuatro partes que mide cuatro propiedades musicales de los acordes.

Aplicando las propiedades del TIS desarrollamos cuatro medidas diferentes que evalúan cuatro propiedades musicales diferentes, a saber, la relación perceptual  $\delta$ , la disonancia  $\xi$ , la relación con la escala  $\lambda$  y la relación armónica  $\phi$ . Estos conceptos musicales se introducirán en esta sección. La información ampliada se puede consultar en el Apéndice A. A continuación, detallaremos cada medida.

### 4.2.1 Relación perceptual $\delta$

En el espacio TIS, hemos codificado una función de minimización que mide la falta de relación perceptual entre el último acorde en la secuencia  $T_{i-1}$  y el acorde candidato  $T_i$ , que se expresa formalmente como una función objetivo (Ecuación 4.1).

$$\delta(T_i, T_{i-1}) = \sqrt{\sum_{k=1}^{M=6} |T_i(k) - T_{i-1}(k)|^2}, \quad (4.1)$$

Obsérvese que la función se basa en la Ecuación 3.2b, que codifica la relación perceptual entre dos configuraciones tonales siguiendo la distancia euclídea.

Esta relación perceptual permite construir una progresión musicalmente coherente. Por lo tanto, estamos buscando generar una progresión de acordes con acordes que comparten algunas notas comunes o están perceptualmente relacionados. Esto significa que intentamos minimizar la distancia entre dos configuraciones tonales. Las distancias pequeñas obtenidas a partir de  $\delta$  se interpretan como configuraciones tonales relacionadas, lo que significa que comparten una serie de propiedades musicales, como notas en común, o una sonoridad similar. Por el contrario, los grandes valores de  $\delta$

representan dos configuraciones tonales no relacionadas, lo que significaría acordes, notas o escalas sin propiedades musicales en común.

### 4.2.2 Disonancia $\xi$

En el espacio geométrico, la propiedad formalizada en la Ecuación 3.3 mide la consonancia perceptual de un sonido, definida como la distancia euclídea al centro del espacio y codificada como una función de maximización. Para preservar la coherencia entre todas las funciones objetivo, tenemos que transformar la maximización en una función de minimización interpretada como función de no consonancia o disonancia. Para ello, restamos un valor normalizado para la consonancia a la unidad 1. Así, la disonancia  $\xi$  puede medirse con la Ecuación 3.3 como

$$\xi(T_i) = 1 - \frac{\|T_i(k)\|}{\sum_{k=1}^{M=6} W(k)}, \quad (4.2)$$

donde  $\|T_i\|$  es la norma de  $T_i$  también interpretada como la distancia desde el origen entre 0 y 1. La distancia euclídea se calcula con todos los componentes del vector TIV. Siguiendo la definición de  $T(k)$ , cada componente está asociado con un peso  $W(k)$  que puede afectar al valor final de disonancia. Por lo tanto, con el fin de normalizar los valores, tenemos que dividir por la cantidad total de pesos  $W(k)$ .

Los valores normalizados de  $\xi$  miden el nivel de disonancia de cualquier configuración tonal o sonido. Una disonancia de una configuración tonal puede ser musicalmente interpretada como una combinación de sonidos que ofenden al oído y necesita ser transformada en consonantes, siguiendo la teoría de Rameau [Lester, 2002]. El concepto de disonancia siempre ha jugado un papel esencial en la música occidental, y actualmente se aplica aquí para generar una progresión de acordes agradable, lo que significa que buscamos minimizar  $\xi(T_i)$ . Un sonido disonante obtendrá un valor alto de  $\xi$ , y por lo tanto será descartado como un buen candidato para ser usado en la progresión de acordes. Igualmente, si un acorde obtiene valores inferiores de  $\xi$ , el acorde será interpretado como un buen candidato de acuerdo a su bajo nivel de disonancia.

### 4.2.3 Relación con la Escala $\lambda$

La relación con la escala o la tonalidad  $\lambda(T_i, T_l)$  representa una función objetivo que mide la relación entre el vector TIV de la escala  $T_l$ , que contiene todas las notas de la escala correspondiente a la tonalidad, y el vector TIV del acorde candidato  $T_i$ . La relación con la escala  $\lambda$  se formaliza con la Ecuación 3.2a y se usa para estimar el ángulo entre ambas configuraciones tonales de acuerdo a:

$$\lambda(T_i, T_l) = 1 - \cos \theta = 1 - \frac{T_i \cdot T_l}{\|T_i\| \|T_l\|}. \quad (4.3)$$

La escala define perceptualmente una jerarquía de alto nivel que regula muchos elementos de nivel inferior, incluyendo algunos acordes y notas de una composición, y descartando otras configuraciones. Los criterios para seleccionar un acorde como un buen candidato dependen del número de características que comparten con la escala definida por la tonalidad. En el espacio TIS, estos acordes estarán cerca de la escala, lo que significa que el ángulo  $\lambda$  debe ser mínimo para asegurar que  $T_i$  está *alineado* con  $T_l$ .

La Figura 4.1 (a) ilustra el cálculo de  $\lambda$ , mostrando  $T_i$ ,  $T_l$ , y el ángulo  $\theta$  entre ellos. Entonces,  $\lambda$  varía entre 0 y 2 cuando  $\theta$  varía entre  $\pi$  y 0. Como estamos calculando el ángulo y no la distancia euclídea,  $\lambda$  sólo considera la relación entre el acorde candidato y la tonalidad, descartando la disonancia  $\xi(T_i)$ .

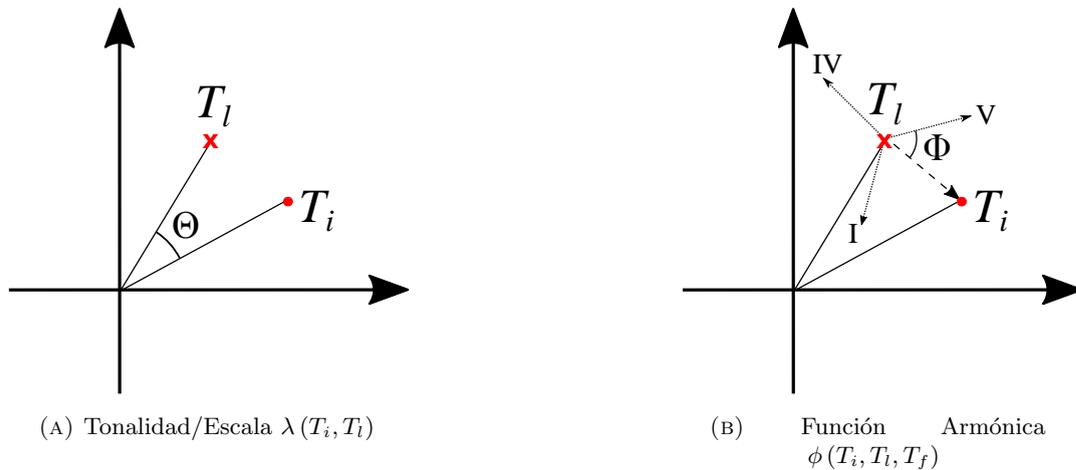


FIGURA 4.1: Interpretación vectorial del cálculo de  $\lambda$  para la tonalidad y  $\phi$  para la función armónica. (a) Cálculo de  $\lambda$  utilizando el ángulo  $\theta$  entre el acorde candidato  $T_i$  y el centro tonal  $T_l$ . (b) Cálculo de  $\phi$  utilizando el ángulo  $\Phi$  entre el acorde candidato  $T_i$  y las funciones armónicas I, IV, y V ambos referenciados por  $T_l$ .

#### 4.2.4 Relación con la Función Armónica $\phi$

La relación con las funciones armónicas  $\phi$  representa la afinidad entre el acorde candidato  $T_i$  y la función armónica  $T_f$  y se codifica como una función de minimización de acuerdo a la Ecuación 4.4.

$$\phi(T_i, T_l, T_f) = 1 - \cos \Phi = 1 - \frac{(T_i - T_l) \cdot T_f}{\|(T_i - T_l)\| \|T_f\|} \quad \text{with } T_f = [I, IV, V], \quad (4.4)$$

Las funciones armónicas determinan la estructura armónica de una composición musical y están siempre relacionadas con la tonalidad. En consecuencia, las funciones armónicas y el acorde candidato utilizan la codificación de la escala  $T_l$  como referencia para calcular la relación entre el acorde  $T_i$  y las funciones armónicas. Hemos incluido esta referencia calculando  $T_i - T_l$  para el acorde candidato y  $T_f$ , que es un vector que representa una de las funciones armónicas I, IV, y V también referenciado por  $T_l$ . La Figura 4.1 (b) ilustra el proceso en el cual tanto los acordes candidatos  $T_i$  como las funciones armónicas son referenciados por  $T_l$  antes de calcular la relación  $\phi$  entre ellos.

Las funciones armónicas se clasifican comúnmente en tres grupos: Tónica (I), Subdominante (IV) y Dominante (V) [Riemann, 1893], que definen juntos la estructura principal de la armonía. Los números asociados a las funciones armónicas están relacionados con su posición en la escala tonal, considerando que una escala diatónica, que son con las que trabajamos aquí, tiene siete grados, uno por cada nota de la escala. La tónica o primer grado se considera el centro tonal que conduce a la resolución, la dominante es el siguiente en importancia después de la tónica, mientras que el subdominante puede ser considerado como el dominante inferior. En la práctica, las progresiones de acordes tienden a utilizar cíclicamente estas funciones como tónica-subdominante-dominante-tónica y así sucesivamente, para permitir el movimiento musical (para más información, véase Apéndice A). Por lo tanto, intentamos minimizar  $\phi$  usando  $T_f$  alternando cíclicamente a través de las funciones armónicas I, IV y V para asegurar que  $T_i$  esté alineado con uno de ellos.  $T_f$  depende implícitamente de los acordes anteriores  $T_{i-1}$  y  $T_{i-2}$  usados para determinar qué función armónica debe compararse con el acorde actual  $T_i$ . Por lo tanto,  $T_f = T_{i-1, i-2}$  y  $\phi$  también pueden ser codificados como  $\phi(T_i, T_{i-1}, T_{i-2}L)$ .

#### 4.2.5 Función Objetivo $P$

Cada medida descrita en los apartados anteriores, a saber, la relación perceptual  $\delta$ , la disonancia  $\xi$ , la relación con la escala tonal  $\lambda$  y la relación armónica  $\phi$ , puede ser considerada como una función objetivo independiente que debe minimizarse. Si  $\delta$ ,  $\xi$ ,  $\lambda$  y  $\phi$  se combinan linealmente, podemos obtener una función objetivo o de fitness tal y como se muestra en la Ecuación 4.5.

$$P(\delta, \xi, \lambda, \phi) = D\delta + X\xi + L\lambda + F\phi. \quad (4.5)$$

donde  $D$ ,  $X$ ,  $L$  y  $F$  son constantes usadas para ponderar la importancia relativa de cada término. Estos pesos se pueden optimizar o definir empíricamente.

La función objetivo  $P$  es una medida cuantitativa de la calidad de acordes, por lo tanto, minimizar  $P$  significa encontrar acordes de buena calidad para seguir una progresión. Para cada acorde candidato  $T_i$ , la función objetivo  $P(\delta, \xi, \lambda, \phi)$  incluye diferentes medidas cuantitativas que capturan propiedades musicales, como la distancia  $\delta$  entre  $T_i$  y el acorde anterior  $T_{i-1}$ , la disonancia  $\xi$  asociada con  $T_i$ , la relación  $\lambda$  entre  $T_i$  y el centro tonal o la escala  $T_l$  y la relación  $\phi$  entre  $T_i$  y la siguiente función armónica  $T_f$  alternando entre I, IV y V.  $P(\delta, \xi, \lambda, \phi)$  resultará en diferentes valores objetivos para el mismo  $T_i$  dependiendo de la tonalidad  $T_l$  y de los acordes anteriores  $T_{i-1}$  y  $T_{i-2}$ . Hay que tener en cuenta que estos diferentes objetivos codificados como relación perceptual  $\delta$ , disonancia  $\xi$ , relación tonal  $\lambda$  y relación con la función armónica  $\phi$  pueden ser conflictivos. Minimizar  $\xi$  podría aumentar simultáneamente  $\delta$ , resultando en acordes con una peor calidad global (medida por  $P$ ). Matemáticamente, minimizar  $P$  es un problema de optimización multiobjetivo, que implica la optimización simultánea de objetivos múltiples a menudo conflictivos.

### 4.3 Explorando Acordes con un Sistema Inmunológico Artificial (AIS)

En este framework, proponemos integrar un sistema que iterativamente genera una progresión de acordes como un problema de búsqueda en el TIS [Bernardes et al., 2016] sometido a restricciones musicales codificadas en la función objetivo  $P$  que se minimiza. Estas restricciones vendrán dadas por el usuario de antemano, a través de la interfaz del framework. Dada una secuencia de acordes, buscamos el siguiente acorde siguiendo restricciones musicales tales como la relación perceptual  $\delta$ , la disonancia  $\xi$ , la relación con la escala  $\lambda$  y la relación con la función armónica  $\phi$ . En cada iteración, seleccionamos un acorde de una paleta de opciones devueltas por el algoritmo de búsqueda y aplicamos de nuevo el procedimiento de búsqueda iterativo para añadir otro acorde a la progresión. La Figura 4.2 muestra una descripción general del método para cada iteración. Los vectores de croma se convierten en vectores TIV con la DFT aplicando la Ecuación 3.1. La función objetivo a minimizar asigna un valor objetivo a cada vector TIV y el AIS encuentra múltiples mínimos de la función objetivo en paralelo que corresponden a los acordes candidatos para seguir la progresión dada.

Los sistemas inmunológicos artificiales (AIS) son una herramienta muy popular para solucionar problemas de optimización multiobjetivo. En particular, utilizamos opt-aiNet [De Castro and Timmis, 2002a] debido a su capacidad multimodal de encontrar múltiples óptimos de la función objetivo en paralelo. Opt-aiNet es capaz de devolver todos los óptimos de una función multiobjetivo en paralelo dentro de una región de

interés. En el contexto de este trabajo, la capacidad multimodal de opt-aiNet es capaz de proponer varios acordes candidatos de buena calidad para seguir una progresión dada minimizando la función objetivo  $P$ .

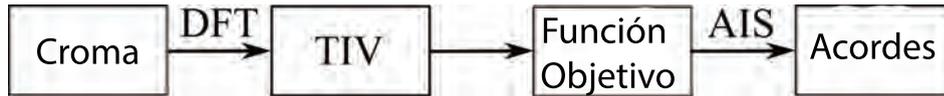


FIGURA 4.2: Representación de la propuesta presentada en este capítulo para generar una serie de acordes candidatos dada una secuencia anterior.

### 4.3.1 Sistemas Inmunes Artificiales y Optimización Multimodal

Los AIS se pueden definir como sistemas adaptativos inspirados por la inmunología teórica y las funciones, principios y modelos inmunes observados, aplicables para resolver problemas en dominios [De Castro and Timmis, 2002a] como la combinatoria [Abdelhaq et al., 2015] o la optimización [Wu et al., 2015]. La optimización multi-objetivo también podría ser un campo muy interesante para aplicar un AIS, ya que típicamente presentan varias posibles soluciones como óptimos de la función objetivo que a veces entran en conflicto. Con el fin de optimizar de manera eficiente este tipo de problemas, opt-aiNet [De Castro and Timmis, 2002a] se desarrolló como una adaptación del sistema inmunológico artificial general. En opt-aiNet, cada célula (anticuerpo) es un vector cuya calidad se mide con una función objetivo. La población de anticuerpos evoluciona siguiendo los principios inmunológicos de expansión clonal, mutación y supresión. La característica más importante de opt-aiNet es su capacidad para buscar el óptimo global y local de la función objetivo de forma paralela. Esto significa que opt-aiNet puede encontrar un conjunto de buenos candidatos para la solución del problema de optimización que son diferentes entre sí.

Opt-aiNet tiene como objetivo optimizar la población midiendo la calidad de los individuos con la función objetivo. Inicialmente, los anticuerpos se inicializan al azar para explorar el espacio de búsqueda. Algunos anticuerpos son seleccionados y clonados en base a su calidad. Cada clon sufre una mutación inversamente proporcional a su valor objetivo, para posteriormente reevaluarlos utilizando nuevamente esta función de optimización. Las células con valores de mayor calidad se introducen en la población. Con el fin de mantener la diversidad, el algoritmo también debe eliminar los anticuerpos que comparten características muy similares. Esta semejanza o afinidad entre dos células puede medirse con una métrica de distancia tal como la distancia euclídea. Los anticuerpos cuya afinidad es superior a un determinado umbral se eliminan de la población manteniendo los que tienen el mayor valor objetivo. Finalmente, una serie de anticuerpos recién generados se incorporan a la red y se repite el proceso.

Idealmente, el procedimiento de búsqueda debe proponer múltiples soluciones candidatas en paralelo, de manera que todas sean buenas opciones y diferentes entre sí. Un mínimo de la función objetivo corresponde a un acorde que es una buena solución candidata para seguir la secuencia actual de acordes. Sin embargo, la función objetivo que se minimiza presenta la mejor solución llamada mínimo global y potencialmente múltiples mínimos en diferentes regiones del espacio de búsqueda, conocidos como mínimos locales. Algunos métodos tradicionales de optimización, como la subida de colina o los métodos del gradiente se pueden quedar atrapados en un mínimo local porque estos procedimientos de búsqueda tienden a encontrar mínimos que están en las proximidades del punto donde la búsqueda se inicializó. La Figura 4.3 ilustra el comportamiento de optimización multimodal de opt-aiNet, mostrando una función objetivo que se minimiza en un intervalo de un espacio de búsqueda unidimensional que va de X a Z. La función objetivo es continua, lo que significa que los puntos cercanos en el espacio de búsqueda tienen valores objetivos similares, dando lugar a variaciones suaves. La función objetivo presenta tres mínimos en el rango de interés en las regiones etiquetadas con X, Y y Z. El objetivo de un algoritmo de optimización que minimiza esta función es encontrar el valor más bajo posible de la función en el intervalo mostrado. Naturalmente, la función puede ser definida para otros valores fuera del intervalo mostrado y puede haber otros mínimos fuera del espacio de búsqueda (posiblemente soluciones aún mejores que las disponibles en el rango de interés). El punto marcado con un triángulo en la región X de la Figura 4.3 es el mínimo global porque es el valor más bajo posible de la función objetivo en el rango mostrado. Los otros dos puntos marcados con un “x” en las regiones Y y Z se llaman mínimos locales porque son los valores más bajos *dentro* de estas respectivas regiones. Los puntos cercanos a cada mínimo marcado con puntos son todos vecinos de los mínimos, y en general representan buenas soluciones al problema en cuestión. Por ejemplo, los puntos  $X_1$  y  $X_2$  en la región X son inferiores a los puntos de la región Y, lo que las convierte en mejores soluciones. Sin embargo, los puntos  $X_1$  y  $X_2$  son vecinos del mínimo marcado con la codificación GM en el espacio de búsqueda, lo que significa que son similares a GM y entre sí, representando soluciones similares. En general, queremos encontrar la mejor solución posible en esa región. Sin embargo, dependiendo de la naturaleza del problema y del rango de interés, los mínimos en Y y Z pueden ser de interés dado que representan la mejor solución posible en esa región del espacio de búsqueda y son diferentes del mínimo en X. Utilizamos opt-aiNet para encontrar múltiples mínimos de la función objetivo (Ecuación 4.5).

La Figura 4.3 (b) ilustra los acordes seleccionados por opt-aiNet para tres regiones de la función objetivo, dada una secuencia de entrada de dos acordes de la escala Do mayor (C mayor), donde se espera un acorde que cumpla con la función de dominante. El mejor candidato es Sol Mayor (GM) en la región X, que es el acorde de dominante (V). La

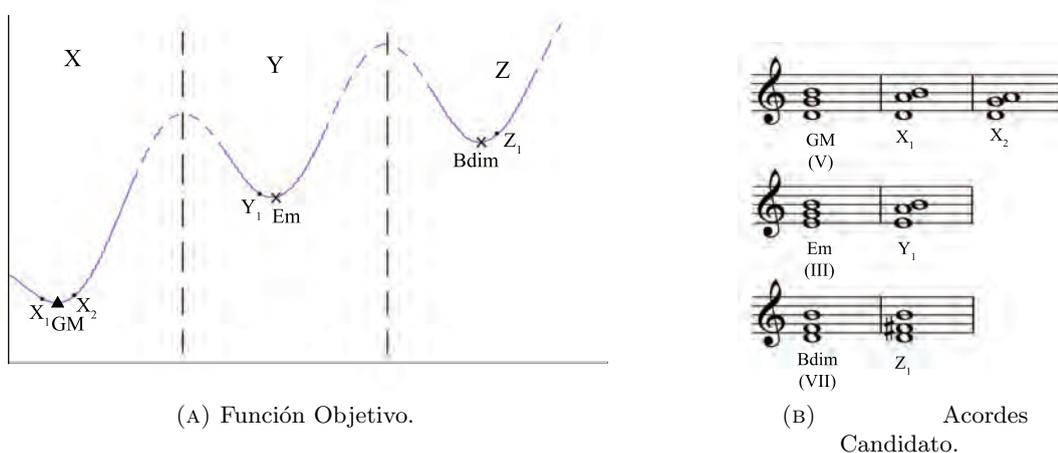


FIGURA 4.3: La figura representa una función objetivo hipotética. El objetivo es encontrar los mínimos correspondientes a los buenos individuos que deben ser preservados en nuestra población. Los puntos resaltados con un círculo son los individuos cerca de los mínimos, los puntos resaltados con una X representan los mínimos locales y el símbolo del triángulo representa el mínimo global. Todos ellos pueden ser considerados buenos candidatos en nuestra población.

figura muestra que el segundo mejor (Mi menor o Em) está en la región Y y el tercero mejor (Si disminuido o Bdim) está en la región Z, lo que sugiere que los óptimos locales corresponden también a soluciones de buena calidad. Los acordes candidatos Em y Bdim corresponden al tono principal (vii<sup>o</sup>) y al mediant (iii), respectivamente. Estos acordes no sólo siguen movimientos típicos entre categorías de funciones en la música tonal, sino que también se ordenan según su relación con el grado dominante esperado. De hecho, el tono principal y estos grados se utilizan a menudo como sustituciones del acorde de dominante.

Además, la Figura 4.3 nos ayuda a entender por qué las medidas de distancia en el espacio TIS o en espacios relacionados no ofrecen un indicador suficientemente bueno de calidad. Los acordes de candidatos que se encuentran en el entorno del mínimo global y local están perceptualmente relacionados con el mínimo local porque comparten tonos comunes y exhiben atributos tales como acordes paralelos (Región Z). Sin embargo, a pesar de su proximidad perceptual, los acordes en el entorno del mínimo local tienen atributos que los descartan de ser buenas soluciones candidatas, tales como acordes altamente disonantes e incompletos (regiones X e Y) e incluyendo notas fuera del conjunto diatónico de la tonalidad. Por lo tanto, los términos en la función objetivo además de la distancia  $\delta$  ayudan a proponer soluciones de buena calidad incluyendo otros criterios musicales, como la disonancia  $\xi$ .

## 4.4 Conclusiones

En este capítulo se ha presentado el diseño de una parte de nuestro framework que compone armonías utilizando un algoritmo evolutivo (Sistema Inmunológico artificial) capaz de encontrar múltiples óptimos simultáneamente, que son diferentes entre sí, lo que se traduce en diferentes acordes candidatos para nuestra composición. Esto es esencial para el objetivo de esta parte del sistema, puesto que cada vez que se obtienen los diferentes candidatos, estos se presentan al usuario, quien selecciona cuál de ellos se incorpora a la progresión de acordes. Esta generación semiautomática facilita la colaboración e integración hombre-máquina, dando como resultado una progresión armónica que se ha adaptado parcialmente al gusto del usuario del sistema. En el Capítulo 7 se describirá más detalladamente esta integración dentro del framework total.



# Capítulo 5

---

## Sociedades Humano-Agente en un Contexto Creativo

---



VNiVERSiDAD  
D SALAMANCA



# Sociedades Humano-Agente en un Contexto Creativo

---

## 5.1 Introducción

Gran parte del comportamiento humano se explica en función de la relación entre los seres humanos y su entorno. La manera de ser o de responder a cualquier situación es diferente para todos, pero estas actitudes están influenciadas por el contexto en el cual la persona fue educada; familia, amigos y experiencias de vida. Estos factores forman la naturaleza social del ser humano [Erickson and Kellogg, 2000].

Como criaturas sociales, miles de decisiones se toman diariamente en base a nuestro entorno social. La información dada por tales contextos sociales proporciona la base para inferencias, planificación y coordinación de cualquier actividad. Sin embargo, este concepto de naturaleza social no puede traducirse en sistemas digitales. En el mundo digital somos socialmente ciegos [Erickson and Kellogg, 2000].

La aparición de la computación social sirvió para resolver este problema e introducir un nuevo concepto: la transparencia social, un enfoque para diseñar sistemas digitales que hace que la información sea visible para los usuarios [Erickson and Kellogg, 2000]. Este nuevo paradigma ha sido aplicado en varios contextos para mejorar un software social para facilitar la interacción y la comunicación entre las personas, para computarizar aspectos de la sociedad humana y para predecir los efectos de las tecnologías en el comportamiento social [Wang, 2007]. En consecuencia, se han creado varias herramientas de servicios web para apoyar comunidades en línea, corregir errores y recopilar comentarios. Muchas herramientas han sido desarrolladas por Microsoft e IBM, HP, Nokia, NASA y Google [Deng and Tavares, 2013, Wang et al., 2007].

Otros autores han calculado modelos de inteligencia social basados en teorías sociales y psicológicas. [Girard et al., 2013, Si et al., 2006, Swartout et al., 2006] han implementado agentes que desarrollan habilidades sociales como liderazgo, lenguas extranjeras y cultura

en una sociedad artificial. Estos modelos sociales enriquecen el comportamiento del agente [Gratch et al., 2006, Pynadath and Marsella, 2005]. Por ejemplo, los Sims 2 [Zaphiris and Ozok, 2012] es un juego popular que modela un mundo virtual con una comunidad social. También podemos considerar robots sociales interactivos tales como Teddy Bear, que fue diseñado por MIT Media Lab [Stiehl et al., 2005] para generar diversas reacciones sociales al tacto humano.

En el área de negocios, las aplicaciones más utilizadas son los sistemas de recomendación para sugerir productos, servicios e información a los consumidores potenciales. Empresas como Amazon o Netflix están adoptando estos sistemas [Wang et al., 2007] para mejorar la lealtad del cliente. Un enfoque es el filtrado colaborativo para predecir las ventas futuras mediante el uso de transacciones de ventas históricas [Huang et al., 2007]. En el sector público, algunas aplicaciones gubernamentales aplican la computación social para detectar organizaciones terroristas, criminales u otras similares [Ferrara et al., 2014, Sánchez et al., 2016]. La computación social también se ha aplicado para apoyar la toma de decisiones en la política de salud o la intervención estatal [Bajo et al., 2015].

En particular, este capítulo se centra en la carga de imágenes por los usuarios para crear una composición musical mediante la traducción de colores en el sonido, basándonos en el fenómeno de sinestesia. Se pretende por tanto, diseñar un sistema descrito como una organización virtual (VO) donde los seres humanos y las máquinas colaboran en un proceso creativo para transformar una imagen en una nube de sonido musical, basándonos en computación social. La VO es implementada por un sistema multiagente (MAS), y define funciones especializadas para extraer sonidos de los píxeles de color de una imagen. Los agentes entonces comienzan un proceso iterativo basado en un algoritmo del enjambre para crear el sonido según diversas reglas sobre inteligencia del enjambre y composición de la música.

La siguiente sección describe una nueva arquitectura para los escenarios de creatividad, basada en la computación social, mientras que la Sección 5.3 detallará la descripción técnica del flujo de trabajo del sistema. Finalmente, en la última sección se analizan las implicaciones de la propuesta.

## 5.2 Arquitectura del Generador de Melodías

Nuestro objetivo es desarrollar un sistema basado en computación social capaz de generar música a partir de imágenes proporcionadas por los usuarios. De acuerdo a la naturaleza de los elementos que intervienen, podemos dividir este sistema en dos componentes: el componente social y el componente virtual o componente máquina. Se

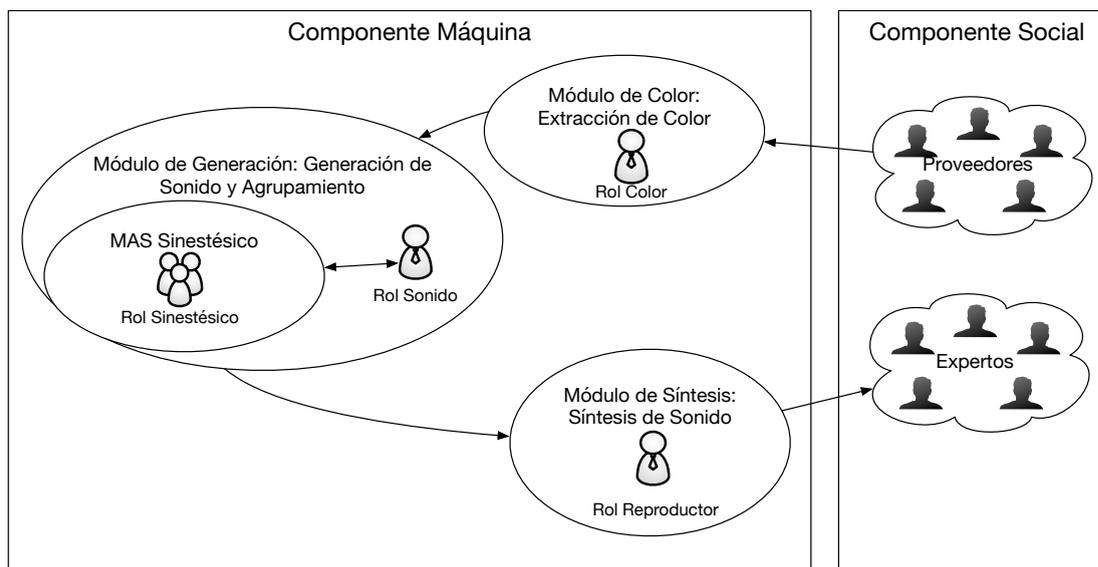


FIGURA 5.1: Esquema del sistema. Las etapas están subrayadas con los recuadros.

proporcionará una descripción introductoria para dar al lector una idea general sobre el sistema desarrollado.

La Figura 5.1 representa un sistema social donde el componente social, seres humanos (proveedores y expertos), colaboran con el componente máquina del sistema inteligente. Los proveedores se centran en proporcionar información sobre las imágenes, mientras que los expertos evalúan la calidad de las composiciones musicales. El enfoque presentado en este trabajo se centra en la arquitectura general del sistema propuesto. El flujo global funciona de la siguiente manera:

- Proveedores. El usuario interactúa con el componente máquina del sistema proporcionando imágenes, esenciales para que nuestro sistema funcione.
- Extracción de color. La primera etapa extrae los colores en la codificación RGB (rojo, verde, azul) de la imagen de entrada. Luego, transforma estos datos RGB en HSB (Tono, Saturación, Brillo).
- Generación de sonido. Transforma el color en notas musicales individuales. Para ello, asocia cada color extraído a un conjunto de notas. Estas notas se agrupan en vecindarios que constituyen la entrada para el paso de agrupación de notas sinestésicas, explicado más abajo.
- Agrupación de notas sinestésicas. Para agrupar notas consonantes, se aplica un algoritmo inteligente en enjambre. Este proceso se basa en el algoritmo de optimización en enjambre de partículas (PSO) [Kennedy, 2010], donde cada elemento (nota) se considera una partícula con fuerzas de atracción o de repulsión.

Estas fuerzas se establecen siguiendo una función de aptitud que evalúa la calidad del sonido producido por cada nota y sus vecinos.

- Generación de la melodía. El paso final determina qué sonidos se seleccionan como un sonido en la composición actual, así como el orden en que los sonidos se reproducirán de acuerdo con diferentes parámetros. Por último, la información MIDI se transforma en sonido físico.
- Expertos. El usuario recibe el sonido físico y lo evalúa. Esta evaluación se detallará en el Capítulo 7.

La siguiente sección describirá la arquitectura general de la VO utilizada para desarrollar el sistema.

### 5.2.1 Arquitectura Virtual

Los diferentes pasos descritos anteriormente deben implementarse utilizando diferentes módulos inteligentes. Esta implementación puede ser intercambiable dependiendo del contexto. Por ejemplo, el módulo correspondiente al sistema de enjambre podría ser sustituido por un módulo inteligente basado en un modelo estadístico (por ejemplo, modelos de Markov) que mueve las partículas de acuerdo con sus probabilidades para aparecer en una composición musical considerando un estilo de música específico. Además, se puede implementar un algoritmo diferente para seleccionar el orden de las partículas en la nube de sonido, aplicando una búsqueda espacial basada en diferentes medidas matemáticas entre partículas. Con un MAS no sólo es posible hacer un sistema distribuido, sino integrar fácilmente componentes inteligentes, combinando incluso diferentes tecnologías o lenguajes.

El análisis y diseño de tal arquitectura abierta requiere un proceso específico dado por organizaciones virtuales (VO). Esta opción permite volver a implementar, cambiar o mejorar algunos módulos en un trabajo futuro. Para ello, hay que definir los principales conceptos desarrollados en la VO, tales como roles, reglas, grupos o funcionalidad de comunicación. GORMAS fue seleccionado como una metodología clave para describir el sistema. Como se detalla en el Capítulo 2, GORMAS es una metodología basada en cuatro dimensiones (estructurales, funcionales, ambientales y normativas) que permiten definir las organizaciones de agentes. El resultado de cada dimensión es uno o varios esquemas globales denominados artefactos. Aunque GORMAS implica un extenso estudio que produce muchos esquemas, por brevedad, aquí solo se presentan dos artefactos o esquemas obtenidos durante las diferentes fases y perspectivas organizativas propuestas por GORMAS.

Durante la primera etapa, al desarrollar la definición del sistema utilizado por la VO, es necesario analizar las necesidades y expectativas de los usuarios potenciales del sistema. El resultado de este análisis es el conjunto de roles involucrados en el sistema propuesto. La Figura 6.1 muestra las funciones principales identificadas en cada módulo descrito anteriormente.

- Proveedor: Representa la primera parte del sistema social. En este caso, el usuario será el proveedor de la imagen de entrada y el oyente del resultado final.
- Color: Extrae los colores de la imagen que se asociará a los agentes Sinestésicos.
- Rol Sinestésico: Transforma el color en sonido. Para ello, cada color está asociado a un agente Sinestésico. Todos los agentes sinestésicos son partículas sumergidas en un algoritmo swarm que les permite navegar por el espacio y cambiar su conocimiento sobre los sonidos a medida que se mueven. Al final, los diferentes agentes se agruparán según su afinidad con respecto a los aspectos musicales que se describirán en la Sección 5.3.3.
- Sonido: Decide el orden en el que los grupos de sonidos correspondientes a grupos de agentes sinestésicos se reproducirán de acuerdo con diferentes parámetros como consonancia o la afinidad perceptual.
- Reproductor: Transforma las notas numéricas en información MIDI que puede generar música física y crea la salida de música basada en restricciones MIDI.
- Supervisor: El supervisor es un agente común en cada organización virtual. Un agente que ejerza esta función tendrá un control general del sistema. Analiza la estructura y la sintaxis de todos los mensajes dentro y fuera del sistema. Como es un agente técnico no relacionado con el trabajo principal, por claridad no está representado en la Figura 6.3.

Aplicamos la metodología GORMAS en el proceso de diseño de este módulo generador de melodías. La Vista funcional mostrada en la Figura 5.2 es una de las vistas esenciales, ya que incluye los recursos consumidos al producir los servicios ofrecidos para lograr las misiones del sistema. Esta organización realiza cuatro misiones, a saber: (i) maximizar la aceptación social de la música generada, (ii) maximizar la eficiencia en la transformación del color en música, (iii) maximizar la eficiencia en la síntesis de los datos matemáticos y (iv) maximizar el impacto de la música y las interacciones sociales en el sistema.

Para completar estas misiones, el sistema consume un servicio, la imagen de entrada específica proporcionada por los usuarios. Los elementos ofrecidos son el producto final

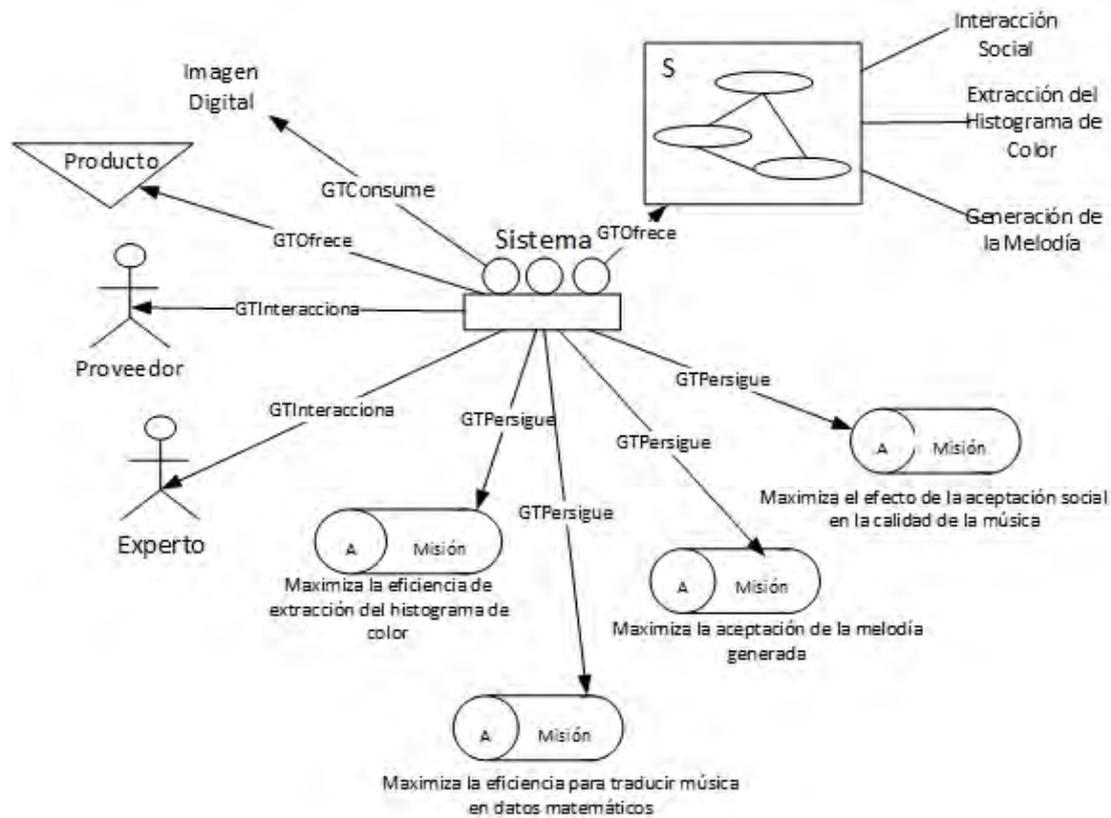


FIGURA 5.2: Vista funcional siguiendo la metodología GORMAS. Destacamos las misiones que el sistema debe llevar a cabo.

y cuatro servicios: síntesis de sonido, generación de música, transformación de color e interacción social.

La Sección 5.3 explica la implementación concreta de cada rol diseñado en la VO.

### 5.3 Descripción del Sistema

Esta sección detalla el flujo de trabajo del sistema que contiene los algoritmos y técnicas implementadas. Se divide en cuatro subsecciones que describen las cuatro etapas del sistema. La Sección 5.3.1 presenta la interacción entre usuario y máquina. La Sección 5.3.2 detalla la extracción de color de la imagen proporcionada por los usuarios y la generación de los agentes sinestésicos, lo cual es esencial para crear música. La Sección 5.3.3 y 5.3.4 explican el modelo de interacción entre los agentes sinestésicos. La Sección 5.3.5 describe un agente de sonido que agrupa y ordena el sonido creado por los agentes sinestésicos mientras que la Sección 5.3.6 presenta el proceso de síntesis para reproducir música, llevado a cabo por el Agente de Reproducción. Finalmente, la Sección 5.3.7 detalla brevemente el papel de los usuarios como evaluadores del producto obtenido por el sistema.

### 5.3.1 Interacción con el Proveedor

El proceso se inicia cuando un usuario carga una imagen en el sistema. Este paso es esencial para iniciar todo el sistema. Esta tarea se puede realizar automáticamente si el sistema selecciona una imagen aleatoria de Internet. Sin embargo, los humanos generalmente no seleccionan las imágenes de forma aleatoria, sino que eligen archivos digitales de acuerdo con sus gustos personales o porque tienen curiosidad por la composición musical de una imagen específica. En consecuencia, estamos más interesados en la interacción social entre un ser humano que intencionalmente selecciona una imagen y la computadora. Por lo tanto, este paso sólo puede ser llevado a cabo por los usuarios humanos.

### 5.3.2 Extracción del Color

El proceso de extracción del color consiste en recibir la imagen digital que el humano proporciona y extraer los colores de la imagen. En el primer paso, el agente Color crea una cuadrícula de celdas como se muestra en la Figura 5.3. El número de celdas en la cuadrícula se establece de antemano por el usuario y debe ser un número entero inferior al número de píxeles de la imagen. El color de cada celda corresponderá al color medio entre todos los colores existentes en el área estudiada. Entonces, el RGB se transforma en HSB. Estas tres características se utilizan para transformar el color en sonido para instanciar los agentes sinestésicos (Figura 5.3). Para aclarar este proceso, el color RGB resulta de una combinación entre varios colores Rojo, Azul y Verde entre 0 y 255. HSB resulta de una combinación de una cantidad de Tono entre 0 y 360, Saturación de 0 a 100 y Brillo de 0 a 100. Por ejemplo, un color verde en RGB se representa como un punto  $\langle 0, 255, 0 \rangle$  y  $\langle 120, 100, 100 \rangle$  en HSB. Su traducción se puede hacer fácilmente con diferentes algoritmos incluidos en algunas aplicaciones multimedia, como Photoshop.

La siguiente sección describe el proceso y la implementación de la función sinestésica.

### 5.3.3 Generación de Sonido

Una vez que se ha extraído el color, los agentes sinestésicos se colocan en un espacio 2D, como podemos ver en la Figura 5.3. En la primera etapa, la posición  $(x, y)$  corresponde al número de la fila y la columna de este cuadrado. Cuando es necesario, estos agentes pueden cambiar sus posiciones codificadas como coordenadas  $(x, y)$ , donde  $x$  e  $y$  son números reales. Cada agente sinestésico contiene estas propiedades:

- Color: esta propiedad consta de tres atributos siguiendo el modelo HSB.



FIGURA 5.3: El proceso de extracción del color mostrado en esta figura consiste en tres fases: fragmentación, extracción del color principal y dotación del color a los agentes sinestésicos.

- **Posición:** Se considera una posición bidimensional  $(x, y)$  como se ha detallado anteriormente.
- **Sonido:** Son los nombres de las notas en términos de sonoridad y tono como se describió anteriormente.
- **Velocidad:** Una matriz con dos vectores, uno para la velocidad en el eje X y otro para la velocidad del eje Y.
- **Calidad:** Una variable booleana para almacenar la decisión sobre si el sonido es lo suficientemente bueno para ser reproducido.

El color de cada celda (en el modelo HSB), producido por el agente Color, se transforma en un sonido por estos agentes sinestésicos siguiendo un proceso donde las tres propiedades del color (Tono, Saturación y Brillo) son traducidas a propiedades musicales. Estas relaciones pueden llevarse a cabo de diferentes maneras. Se podría utilizar una interacción social para seleccionar el color de cada nota; sin embargo, para el alcance de este estudio, se siguió una relación estándar propuesta por Sanz [2009]. En primer lugar, el Tono de un color se asocia automáticamente al nombre de la nota siguiendo el sistema de Lagresille [Sanz, 2009], que se muestra en la Figura 5.4 (a), donde cada conjunto de tonos de color corresponde a una nota específica. Por su parte, la Saturación está relacionada con la intensidad del sonido. Podemos considerar que

esta asociación es lógica, ya que cuanto más intenso es el color, más intenso debería ser el sonido. Finalmente, el brillo se refiere al tono real. A medida que aumenta el valor de brillo, el sonido tiene un tono más alto y, por lo tanto, será más agudo. La Figura 5.4 (b) ilustra el proceso de asociación considerando las tres medidas para el sistema HSB. Las flechas muestran el aumento de los valores de brillo y saturación.

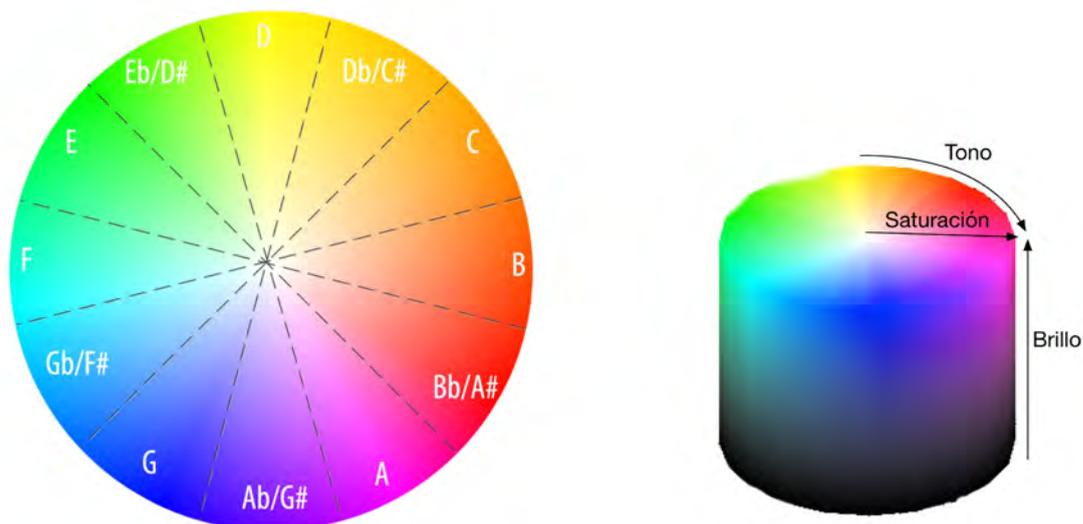


FIGURA 5.4: Figura a: Representación cromática de los tonos y colores. Cada área delimitada por las líneas de puntos pertenece a una nota musical siguiendo el sistema de Lagresille. Figura b: Asociación entre las propiedades de los colores y el sonido siguiendo la asociación propuesta entre el color y el sonido.

En conclusión, los agentes sinestésicos forman una comunidad de agentes, que es el corazón del sistema y donde se desarrolla la transformación real color-musical.

### 5.3.4 Agrupación de Notas

El comportamiento de los agentes que implementa el rol sinestésico se basa en un algoritmo de optimización de enjambre de partículas (PSO) [Kennedy, 2010]. Por lo tanto, el movimiento se regula por fuerzas de atracción capaces de modificar su posición y velocidad después de una función de aptitud. El enjambre permite la asociación de varios agentes con características similares después de una función de aptitud estudiada en la Sección 5.3.4.1. Esta subsección describe el proceso de enjambre y los resultados obtenidos resumidos en el Pseudocódigo 2.

Los pasos seguidos en este algoritmo son:

1. Cada agente tiene una posición  $P$  en el espacio 2D, y puede producir un sonido asociado al color que representa.

**Algorithm 2** Proceso en Enjambre**Algorithm 5.3.1:** AGENTESINESTESICO()

---

```

for each  $a \in \mathcal{A}$ 
  do  $\left\{ \begin{array}{l} a.position \leftarrow \text{EXTRAERPOSICIONDECOLOR}(a.color) \\ a.sound \leftarrow \text{EXTRAERSONIDO}(a.color) \end{array} \right.$ 
  for each  $a \in \mathcal{A}$ 
    while  $PosicionAnt - a.position > UMBRAL$ 
      do  $\left\{ \begin{array}{l} Vecinos \leftarrow \text{EXTRAERVECINOS}(a) \\ \textbf{for each } n \in Vecinos \\ \textbf{do } \left\{ \begin{array}{l} fitTemp \leftarrow \text{MEDIRFITNESS}(a.sonido, n.sonido) \\ fuerzaTemp \leftarrow fuerzaTemp + \text{GENERARFUERZA}(fitTemp) \end{array} \right. \\ posicionAnt \leftarrow a.position \\ a.position \leftarrow \text{CALCULARPOSICIONFUERZA}(fuerzaTemp) \end{array} \right.$ 

```

---

2. Cada agente  $a_1$  busca en sus vecinos  $a_2, a_3, \dots, a_N$  en el espacio e intercambia información con ellos para medir la consonancia o calidad entre estos sonidos. Este valor provoca una fuerza de atracción entre los agentes. La fuerza depende del nivel de calidad de los sonidos.
3. Cada agente tiene la capacidad de saber si su resultado es lo suficientemente agradable para ser incluido en la melodía. Para ello, se ha definido un umbral para el resultado de la ecuación de consonancia.
4. Estos pasos se repiten durante el tiempo hasta que se encuentra un equilibrio de sonido, después de la convergencia del algoritmo. El balance de sonido significa que las partículas no actualizan sus posiciones significativamente sobre las iteraciones. Esto significa que los sonidos están equilibrados en sus posiciones correctas según la función de calidad analizada.

En el estado final, se obtendrá una organización de agentes con agrupaciones de sonidos agradables dentro del mismo grupo.

Como se mencionó anteriormente, cada agente clasifica su calidad según una función de acondicionamiento físico. Esto se describirá en la sección 5.3.4.1. La Sección 5.3.4.2 está dedicada al modelo de interacción sinestésica implementado por la VO.

#### 5.3.4.1 Un Valor de Aptitud para Medir la Calidad del Sonido

Anteriormente se indicó que la función de aptitud afecta a la intensidad de las fuerzas de atracción entre los agentes. Para modelar la función de aptitud o función de fitness,

se han considerado dos factores musicales: la consonancia y la sonoridad.

Para hallar la función de fitness o aptitud del PSO nos hemos basado en dos conceptos musicales que son clave para obtener una secuencia agradable y consonante de sonidos: la intensidad y la afinidad sonora (para más información sobre la teoría musical, véase el Apéndice A). Para poder trabajar con estos conceptos, hemos analizado los sonidos en el Tonal Interval Space (TIS), tal y como se detalla en el Capítulo 2.

Una vez los sonidos están representados en este espacio, podemos analizar su afinidad sonora a partir de dos conceptos que nos permiten hallar el valor de la función de fitness de las partículas:

- Consonancia. Para hallar la consonancia de un sonido, se ha calculado la distancia euclídea de esta frecuencia al centro (punto 0 en el espacio).
- Relación perceptual entre notas. Este concepto musical se ha calculado, físicamente, como la distancia euclídea entre las frecuencias de dos notas.

Otro factor a tener en cuenta tiene relación con el volumen de la música. Es deseable equilibrar el volumen para notas consecutivas (minimizar la distancia entre dos niveles de sonoridad consecutivos). De lo contrario, los oyentes encontrarían la música inestable debido a los cambios aleatorios entre el fuerte y el piano.

La función de fitness debe conjugar el término de volumen y de calidad sonora, mediante una combinación lineal de acuerdo a la siguiente ecuación:

$$F(n_a, n) = \sum_{i=0}^M (W_c \cdot C(n_a) W_d \cdot D(n_a, n_i) + W_l \cdot L(n_a, n_i)) \quad (5.1)$$

$$C(n_a) = \xi(T_a) \quad (5.2)$$

$$D(n_a, n_i) = \delta(T_a, T_i) \quad (5.3)$$

Donde cada agente  $a$  contiene un sonido  $n_a$  y un vector  $n$  con los sonidos de los  $M$ -agentes más cercanos.  $C(n_a)$  representa la consonancia del sonido  $n$  del agente  $a$ , mientras que  $D(n_a, n_i)$  mide la relación perceptual entre los dos sonidos.  $T_a$  es la transformada DFT del sonido  $n_a$ , y  $T_i$  es la transformada DFT del sonido  $n_i$ . Los valores  $\xi$  y  $\delta$  ya se han estudiado en el Capítulo 4.  $W_c$ ,  $W_d$  y  $W_l$  son constantes empíricamente establecidas en 0.7, 0.4 y 0.25, respectivamente.

Del mismo modo,  $L(n_a, n_i)$  considera el volumen del sonido correspondiente al agente  $a$  y lo compara con el volumen del sonido  $n_i$ .

$$L(n_a, n_i) = \| L(n_a) - L(n_i) \| \quad (5.4)$$

### 5.3.4.2 Integración del Proceso en Enjambre en un MAS

El comportamiento de la organización virtual está inspirado en el comportamiento de un PSO. Este algoritmo propuesto por Kennedy [2010] es un ejemplo de inteligencia de enjambre donde cada individuo se mueve libremente a través del espacio considerando tres factores: el componente de peso de inercia, el componente cognitivo y el componente social. Tal y como vimos, la fuerza de inercia está relacionada con la inercia física y depende de la fuerza anterior aplicada a la partícula; los componentes cognitivos se refieren a las fuerzas de atracción entre partículas o grupos de partículas y, finalmente, el componente social está relacionado con el intercambio de información entre partículas.

Por tanto, la posición siguiente  $p_t$  de una partícula  $a$  en el enjambre depende de la posición actual  $p_{t-1}$ , la velocidad actual  $v_{t-1}$ , la mejor posición en la hora actual  $pb_i$ , y la posición óptima encontrada por cualquiera de sus vecinos  $pb_n$ , siguiendo la Ecuación [Kennedy, 2010]:

$$p_t = f(p_{t-1}, v_{t-1}, pb_i, pb_n) \quad (5.5)$$

El algoritmo PSO necesita ser adaptado para resolver nuestro problema específico. Las partículas en el algoritmo están representadas por los agentes sinestésicos. Los tres factores que provocan el movimiento de las partículas también deben adaptarse a nuestro sistema creativo. Así, las fuerzas de atracción están relacionadas con la inercia y los componentes cognitivos, mientras que el intercambio de información de los agentes es el componente social. La actitud cooperativa es también esencial en la organización virtual para lograr el objetivo de todo el sistema.

En este caso, la comunicación permite al agente conocer los colores, sonidos y posición de sus vecinos. Estos agentes pueden tener un comportamiento cooperativo y no cooperativo. Las interacciones cooperativas se basan en una función de atracción. La intensidad de las fuerzas de atracción depende de una función llamada función de aptitud  $F(x, y)$ , formalizada en la Ecuación 5.1. Esta función permite modificar la posición de cada partícula según las medidas de consonancia. Por el contrario, las interacciones no cooperativas se refieren a una función de repulsión. Esta función de repulsión se activa sólo si las posiciones de los agentes están muy próximas entre sí, para evitar colisiones

siguiendo la teoría de Blackwell [2003]. En nuestro caso, si se produce un umbral de distancia mínima entre dos partículas, se aplica un vector de velocidad  $v$  a cada partícula  $a$  con dirección (dirección opuesta al vector de velocidad de  $a$ ) y módulo (el módulo de la velocidad es  $V_a + 0,2$ )

El algoritmo comienza cuando cada agente busca a sus vecinos. Para ello, se establece una relación para que el agente seleccione quiénes son sus vecinos. Empíricamente, fijamos este umbral a 1. Los agentes llevan a cabo un proceso de interacción para intercambiar información, y finalmente decidir la mejor posición según la fuerza de atracción generada. La fuerza para el Agente  $a_i$  depende de los valores obtenidos aplicando la función de aptitud según su vecindario, pero esta puede no ser la mejor posición encontrada por el enjambre entero. En consecuencia, este enfoque se denomina enfoque local mejor.

La Ecuación 5.6 representa los cálculos para obtener la siguiente posición  $p_{t+1}$  de un agente dado  $a_i$  en la iteración  $t$ . Esta es una combinación lineal de la velocidad actual  $v_{t+1}$  y la posición anterior  $p_t$ .

$$p_{t+1} = p_t + v_{t+1} \quad (5.6)$$

Hay que tener en cuenta que todos los valores anteriores son vectores. La velocidad  $a_i$  de la partícula viene dada por la Ecuación 5.7.

$$v_{t+1} = \sum_{k=1}^N F * v_t + (p_{kt} - p_t) \quad (5.7)$$

donde  $k$  representa el  $k$ -agente vecino presente para el agente  $a_i$ . Los tres componentes diferentes descritos anteriormente (inercia, cognitiva y social) están representados cada uno por uno de los tres términos de la ecuación 5.7. La función de aptitud  $F$  regula el efecto del componente de fuerza (velocidad). El vector  $(p_{kt} - p_t)$  permite el movimiento de las partículas hacia la mejor posición encontrada por todos los agentes vecinos.

Dentro de cada iteración, cada partícula se mueve en una dirección determinada por la influencia que sus vecinos tienen sobre ella. En nuestro caso, a diferencia del algoritmo PSO general, no existe una mejor posición global para todo el sistema en los pasos intermedios. Las partículas se mueven alrededor del espacio de búsqueda basado en estas ecuaciones durante una serie de iteraciones hasta que, si todo va bien, todos convergen. El criterio de convergencia se logra cuando las posiciones de las partículas no se modifican notablemente. El mejor global entonces puede tomarse como la solución final producida por el algoritmo. En el punto final del algoritmo, también esperamos que se generen

diversos subgrupos con similares características musicales, debido al efecto de las fuerzas de atracción.

Para reproducir la composición completa, cada agente tiene la capacidad de decidir si debe ser reproducido, en otras palabras, modificar la propiedad de “calidad”, que es un valor booleano. Este valor puede ser modificado de acuerdo a si el sonido es suficientemente bueno en términos de consonancia con los vecinos. Para lograr esto, se establecen dos umbrales para que, si los valores para la función de aptitud no están por encima de estos valores de umbral, no sean candidatos para crear la melodía por el Rol Sonido (descrito más adelante). En consecuencia, la propiedad “calidad” se establece en 0. De lo contrario, la propiedad “calidad” se establecerá en 1.

### 5.3.5 Generación de la Melodía

Una vez agrupadas los agentes sinestésicos por consonancia, hay que ordenar sus sonidos para obtener la melodía, donde cada sonido tendrá un determinado tono, volumen y duración. Existirán algunos sonidos que no podrán utilizarse debido a su incompatibilidad dentro del contexto melódico. Para decidir qué agentes se seleccionan es necesario obtener la escala musical que define el contexto tonal de una composición musical. La búsqueda de la tonalidad se ha llevado a cabo siguiendo el modelo de Krumhansl, que permite obtener jerarquías tonales mediante el uso de perfiles tonales basados en los armónicos acústicos de las diferentes notas musicales [Krumhansl and Kessler, 1982].

Posteriormente, se ha elegido la primera nota de la melodía, que será el primer agente del enjambre cuyo sonido coincida con la tónica. Cada vez que una nota se añade a dicha melodía, se utiliza uno de los atributos booleanos del agente para indicar que ya ha sonado, y de esta manera, evitamos repeticiones.

Es necesario seguir un protocolo para añadir una nueva nota en la melodía, una vez seleccionada la primera nota, configurando las propiedades de tono, volumen y duración. En esta sección se estudiará la compatibilidad en cuanto a la propiedad de tono, y luego se decidirá la duración de la nota en función de los parámetros obtenidos.

El PSO nos ha permitido aproximar agentes cuyas propiedades sonoras son afines entre sí. De este modo, la nota a elegir como nota siguiente estará entre los agentes vecinos de la última nota seleccionada.

Puede ocurrir que para la última nota añadida a la melodía no exista ningún agente vecino compatible con dicha nota. En ese caso, se aumenta el radio de búsqueda de vecinos para continuar con la composición. Si, después de buscar en todo el espacio, no

se ha encontrado un agente vecino compatible con la última nota, la generación de la melodía finaliza.

Otra casuística es que existan varios vecinos cercanos a la última nota seleccionada. En ese caso, se analizará cuál de los sonidos es más compatible para añadir a la melodía considerando tres parámetros de calidad:

- **Compatibilidad con la tonalidad.** Para calcular la compatibilidad tonal, calculamos el ángulo  $\Theta$  entre la representación vectorial de la nota en el TIS  $T_i$  y la representación vectorial de la tonalidad  $n_l$  (mayor o menor) en dicho espacio  $T_l$ , tal como está descrito en el Capítulo 3 .

$$\Theta(n_i, n_l) = \theta(T_i, T_l) \quad (5.8)$$

Cuanto más cercanas se encuentren en el espacio, más compatibles serán entre ellas.

- **Atracción melódica.** La atracción melódica recoge la compatibilidad entre dos notas consecutivas teniendo en cuenta su similitud sonora y el número de semitonos que las separan, de acuerdo a la ecuación:

$$A(n_{i-1}, n_i) = D(n_{i-1}, n_i) * e^{w \cdot n_{Sem}} \quad (5.9)$$

El valor del peso del número de semitonos,  $w$ , es 0.2 y ha sido empíricamente obtenido.

- **Afinidad.** La afinidad se calculará siguiendo el modelo detallado en la sección anterior.

Así, la compatibilidad  $\kappa$  es la combinación lineal de la compatibilidad tonal, la atracción melódica y la afinidad.

$$\kappa(n_i, n_{i-1}, n_l) = \Theta(n_i, n_l) + W_{at} \cdot A(n_{i-1}, n_i) + W_c \cdot F(n_{i-1}, n_i) \quad (5.10)$$

Los valores de los pesos también han sido hallados experimentalmente para conseguir una mayor compatibilidad entre las notas y, por tanto, un mejor resultado sonoro. Son los siguientes:  $W_t = 1.14$ ,  $W_{at} = 7$ ,  $W_c = 1.3$ .

La asignación de una duración concreta a una nota de un agente se realiza a partir del valor de la compatibilidad entre la nota anterior y ella misma y la distancia en el espacio bidimensional. Los valores de la compatibilidad están comprendidos en un intervalo de

4 a 11. Las notas con mejor compatibilidad y menor distancia tienen un valor menor, y las menos compatibles, un valor mayor.

En este estudio preliminar, se han utilizado cuatro figuras rítmicas, desde la corchea, que dura medio pulso, hasta la redonda, que dura 4 pulsos. Para mapear las duraciones de las notas, se calcula la media entre la distancia normalizada y la compatibilidad, de manera que si el valor es inferior a 0.25, entonces la duración es una corchea, si es inferior a 0.5 la duración es una negra, con valores hasta 0.75, la duración es una blanca, mientras que para el resto de casos, la duración sería una redonda.

### 5.3.6 Roles de Síntesis y Reproducción

A medida que avanzamos en esta sección, el último paso de nuestra VO tiene como objetivo sintetizar los resultados propuestos por el algoritmo anterior. Los números de las notas, la octava y la sonoridad obtenida necesitan ser interpretados para que un instrumento o un sintetizador pueda reproducirlos. Se considera que el formato MIDI transforma las propiedades de los agentes en datos MIDI [Hewlett and Selfridge-Field, 1997]. Esta es la tarea principal del papel del Agente Reprodutor. Una vez realizada esta tarea, el agente de funciones extrae la información MIDI y la transforma en información de audio para que la computadora pueda reproducirla.

### 5.3.7 Evaluación Experta

En general, los procesos creativos y los productos son difíciles de validar, y ninguna máquina es capaz de hacerlo. En consecuencia, esta tarea sólo puede ser realizada por seres humanos. Es esencial considerar que la evaluación musical depende de cada persona, de su cultura, de su educación musical o de sus preferencias. Por lo tanto, necesitamos una gran cantidad de individuos para validar tales datos subjetivos y extraer conclusiones sólidas.

El proceso creativo y los productos son difíciles de validar, por lo que varios experimentos musicales han sido evaluados siguiendo una evaluación de expertos. Algunos ejemplos serían [Collins et al., 2016, Delgado et al., 2009, Herremans et al., 2015, Navarro et al., 2015], entre otros.

En este caso particular, la salida generada por el componente máquina es evaluada por la interacción experta. La evaluación presentada en este trabajo tiene como objetivo investigar si los acordes con valores de aptitud bajos son juzgados más consonantes que acordes con valores de penalización más altos; en otras palabras, si la función de fitness mide la aceptación social de la música generada.

La evaluación consiste en clasificar cada fragmento generado en una escala de 1 (muy malo) a 5 (muy bueno). La comunidad social puede aplicar esta forma para estudiar la calidad de las composiciones extraídas por el sistema. Esta validación se detallará en el Capítulo 7

## 5.4 Conclusiones

En este capítulo se ha desarrollado un sistema basado en computación social para transformar colores en música. Este sistema contiene dos componentes: social y computacional. El componente social permite tanto proveer al sistema con los datos de entrada (imágenes digitales) como evaluar los resultados finales. El componente máquina contiene un flujo de trabajo de cuatro etapas: extracción del color de la imagen, generación del sistema en enjambre a partir de los agentes sinestésicos, la ordenación de sonidos para la generación melódica y la sintetización y posterior reproducción del resultado final. Todos ellos están basados en organizaciones virtuales e implementados por un sistema multiagente.

Con el fin de evaluar la calidad de los sonidos según se van generando, se ha diseñado una función matemática basada en el espacio TIS presentado en el Capítulo 3, que evalúa la consonancia, la distancia entre notas y la atracción tonal. Hay que tener en cuenta que, aunque una de las fases planteada es recoger la evaluación de los expertos, esta opinión no afecta el resultado del sistema. Esta limitación se abordará en el capítulo siguiente para agregar una retroalimentación al sistema con el fin de incorporar automáticamente las evaluaciones de expertos y así mejorar nuestro sistema.



# Capítulo 6

---

## Integrando Aprendizaje Social

---



VNiVERSIDAD  
DSALAMANCA



# Integrando Aprendizaje Social

---

## 6.1 Introducción

Tal como hemos visto en capítulos anteriores, la naturaleza social es inherente a los seres humanos. El contexto social influye en nuestras rutinas diarias, en la cooperación y en la toma de decisiones, a menudo comparándolas con experiencias similares anteriores. La incorporación de tal información compleja en máquinas puede ser un desafío desde un punto de vista técnico. Por lo tanto, diferentes sistemas basados en computación social se han desarrollado para resolver este problema y hace que la información social sea visible para los usuarios [Erickson and Kellogg, 2000].

Ya se ha comentado también que la computación social ha sido ampliamente aplicada en diferentes campos de investigación para mejorar la interacción social entre las personas [Wang, 2007], para apoyar comunidades en línea [Deng and Tavares, 2013], para imitar una comunidad social [Girard et al., 2013] o como sistemas de recomendación [Wang et al., 2007]. Los sistemas sociales pueden ser diseñadas siguiendo varias propuestas. Una posible implementación son las Organizaciones Virtuales (VO). Como hemos visto, las VO tienen algunas propiedades que son útiles para nuestro enfoque, tales como flexibilidad en entornos dinámicos, autonomía o razonamiento.

Estas características pueden ser modeladas usando diferentes metodologías. En particular, el razonamiento basado en casos (CBR) se ha aplicado con éxito para controlar el comportamiento de la organización, permitiendo una mejor adaptación del VO al entorno social basado en la experiencia anterior en la sociedad siguiendo un proceso muy similar a cómo los seres humanos aprenden comportamiento social Corchado et al. [2013].

El presente capítulo describe nuestro sistema basado en computación social ya presentado en el capítulo anterior, que permite la colaboración de seres humanos y máquinas para generar una composición musical que es una línea melódica. Sin embargo, en este caso, la interacción entre diferentes usuarios humanos sobre impresiones

de intereses y resultados musicales se aprovechará para la mejora del propio sistema inteligente. En particular, algunos agentes que forman parte de la VO incorporan una arquitectura de razonamiento basado en casos (CBR) como un mecanismo de inferencia para mejorar y aprender del entorno. El núcleo de aprendizaje de la arquitectura CBR presentado en esta propuesta es el algoritmo de hormigas o Ant Colony Optimization (ACO).

En el capítulo anterior la computación social también se aplicó en un contexto musical para generar música a partir de una imagen digital. El sistema también fue implementada por una VO, y los agentes pertenecientes a la comunidad reprodujeron un comportamiento de enjambre para transformar colores en música coherente. El resultado final es validado por expertos musicales que calificaron los sonidos. Sin embargo, el prototipo no podía aún aprender de la opinión social extraída de la comunidad de expertos. Por lo tanto, ahora hemos enriquecido el componente máquina del sistema mediante la aplicación de una arquitectura CBR para aprender de las experiencias anteriores basadas en la impresión de los expertos y la mejora de la música de acuerdo a los gustos musicales en general. El motor de aprendizaje de la arquitectura CBR se basa en un algoritmo de enjambre estándar, específicamente la optimización de colonia de hormigas (ACO), para encontrar una solución óptima. Los detalles de este algoritmo serán proporcionados en la Sección 6.3.2.

El sistema desarrollado que alberga los escenarios de creatividad, basado en la computación social, se detalla en la Sección 6.2 y las mejoras realizadas se explican en la Sección 6.3. Por último, en la última sección se analizan las implicaciones de la propuesta.

## 6.2 Sistema Social

Nuestro objetivo es crear un sistema basado en computación social y generar música a partir de imágenes digitales. En nuestro capítulo anterior, la implicación humana consistía en cargar imágenes digitales y validar los resultados musicales. La presente propuesta incorpora, además, mecanismos de retroalimentación para considerar la opinión social y así mejorar los resultados musicales. Para ello, aplicamos una arquitectura CBR.

Podemos observar que este sistema contiene dos tipos de miembros, proveedores y expertos. Los proveedores pueden subir una imagen digital cualquiera. En el capítulo anterior se muestra cómo se extrae el color en codificación HSB o HSL. En este capítulo además, los colores extraídos se agrupan para generar un histograma de la imagen

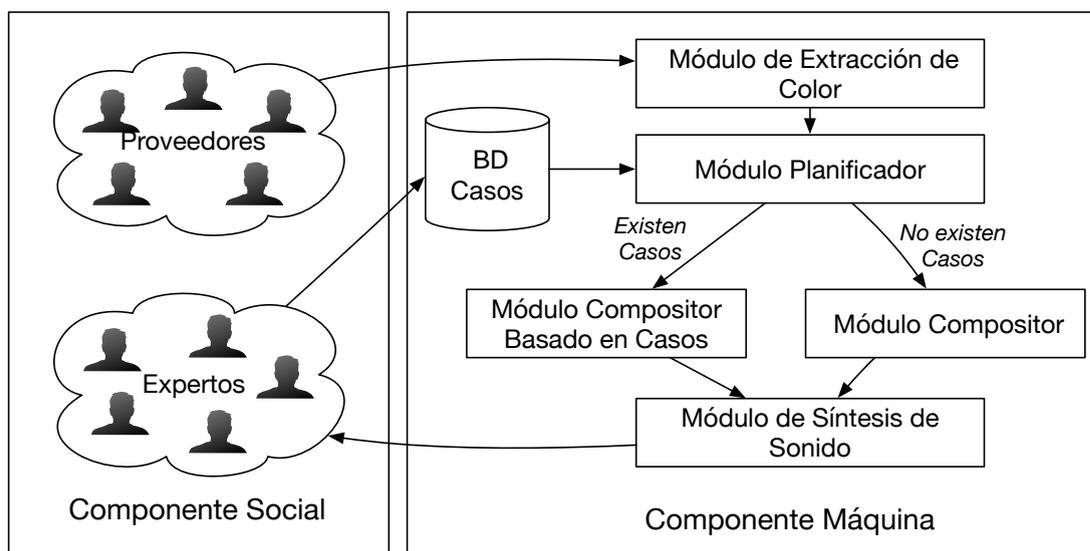


FIGURA 6.1: Flujo de trabajo del sistema. Las etapas se destacan en los cuadros.

como un descriptor con los colores más comunes. El Planificador es capaz de buscar histogramas de color similares en casos anteriores. Estos se utilizan para aplicar un algoritmo ACO y encontrar el mejor camino entre las notas, que supone nuestra nueva solución en el problema. Esta producción musical debe ser evaluada por los expertos. La media de todas las puntuaciones proporcionadas constituye la tasa global de la composición musical. Los expertos pueden tener acceso a las melodías en cualquier momento, por lo que las puntuaciones podrían ser actualizadas o agregadas por nuevos expertos. El Planificador reevalúa el caso cada vez que se recibe una puntuación.

### 6.2.1 Organizaciones Virtuales Aplicadas a Nuestro Problema

Cada etapa de la propuesta aquí detallada puede desarrollarse en diferentes módulos que son intercambiables cuando el contexto es diferente. Una VO nos permite modelar sistemas abiertos que puede combinar módulos inteligentes con diferentes técnicas.

Uno de los pasos esenciales en el diseño del sistema social con VOs, es identificar los roles involucrados en todo el sistema. Estos roles representan las necesidades y expectativas de la organización. En nuestro análisis, se encontraron siete roles, a saber: Proveedor, Color, Planificador, Sonido, Sinestésico, Reproductor y Experto.

El Proveedor da la primera entrada al componente máquina (una imagen digital). En nuestra propuesta esto representa uno de los componentes humanos de nuestro sistema. Una vez que se recibe la imagen, el rol Color extrae el histograma de color. Con este histograma de color, el rol Planificador controla todo el proceso de razonamiento basado en casos (CBR). Este rol es responsable de buscar experiencias anteriores para encontrar

cuáles son similares a la entrada actual y decidir si continuar con el proceso CBR o dar el control a la función compositora. En este último caso, el rol Sonido y el rol Sinestésico genera una nueva melodía, siguiendo el proceso explicado en el capítulo anterior. El resultado obtenido tanto del Planificador como del proceso de composición es sintetizado y reproducido por el rol Reproductor. El resultado musical es evaluado en términos de calidad por el rol Experto, que representa otro componente humano.

Una vez definidos los roles, debemos describir la funcionalidad general con la que queremos dotar al sistema. Esto se puede dividir en misiones o metas a lograr, los recursos necesarios para lograr las metas y los servicios que el sistema puede proporcionar al entorno externo. En este caso, se consideran cuatro misiones: maximizar la exactitud para extraer el histograma de color, maximizar la aceptación social de la melodía generada, maximizar la eficiencia para traducir la información numérica a música y maximizar el efecto de las interacciones sociales. El recurso principal para lograr las misiones es la imagen digital cargada por los proveedores. Además, el sistema puede proporcionar los siguientes servicios: síntesis de sonido, creación de melodía, extracción de histograma de color e interacción social. Esta vista de funcionalidad descrita también se representa en la Figura 6.2.

Para clarificar el funcionamiento general del sistema, se ha añadido la Figura 6.3, donde cada número indicadas diferentes etapas por las que debe pasar el sistema. A continuación, se describirá un ejemplo.

Inicialmente y como parte de la interacción humana, un usuario (Proveedor) proporciona una imagen digital, considerada la entrada del componente máquina (Paso 1), con colores en el espectro del rojo, blanco y azul. A continuación, el Agente Color extrae el histograma de color de la imagen (paso 2) con los colores más comunes de la imagen, en nuestro caso pueden ser rojo, rosa, azul y blanco. El histograma de color es analizado por el agente Planificador para buscar casos anteriores (histogramas) con una composición de color en la base de datos (memoria) similar a la del sistema (Paso 3). Puede ocurrir que se hayan encontrado casos similares (Paso 4) con rojo, rosa, azul y blanco como los colores más comunes. Suponemos que se encontraron ocho casos similares con las medidas de similitud asociadas y las calificaciones de expertos, todas ellas con calidad suficiente para ser consideradas como buenos casos. A continuación, estas soluciones anteriores se adaptan y optimizan con un algoritmo Ant Colony Optimization implementado en el propio agente Planificador, que extrae una nueva propuesta (Paso 5b). Si no se hubiera encontrado ningún caso similar adecuado, el control pasa al Agente Sonido y al MAS Sinestésico, responsables de generar música nueva, y así crear un fragmento musical siguiendo el algoritmo de composición propuesto en el capítulo anterior (paso 5a). En ambos casos, se genera una composición musical y se envía al Planificador (pasos 6a y

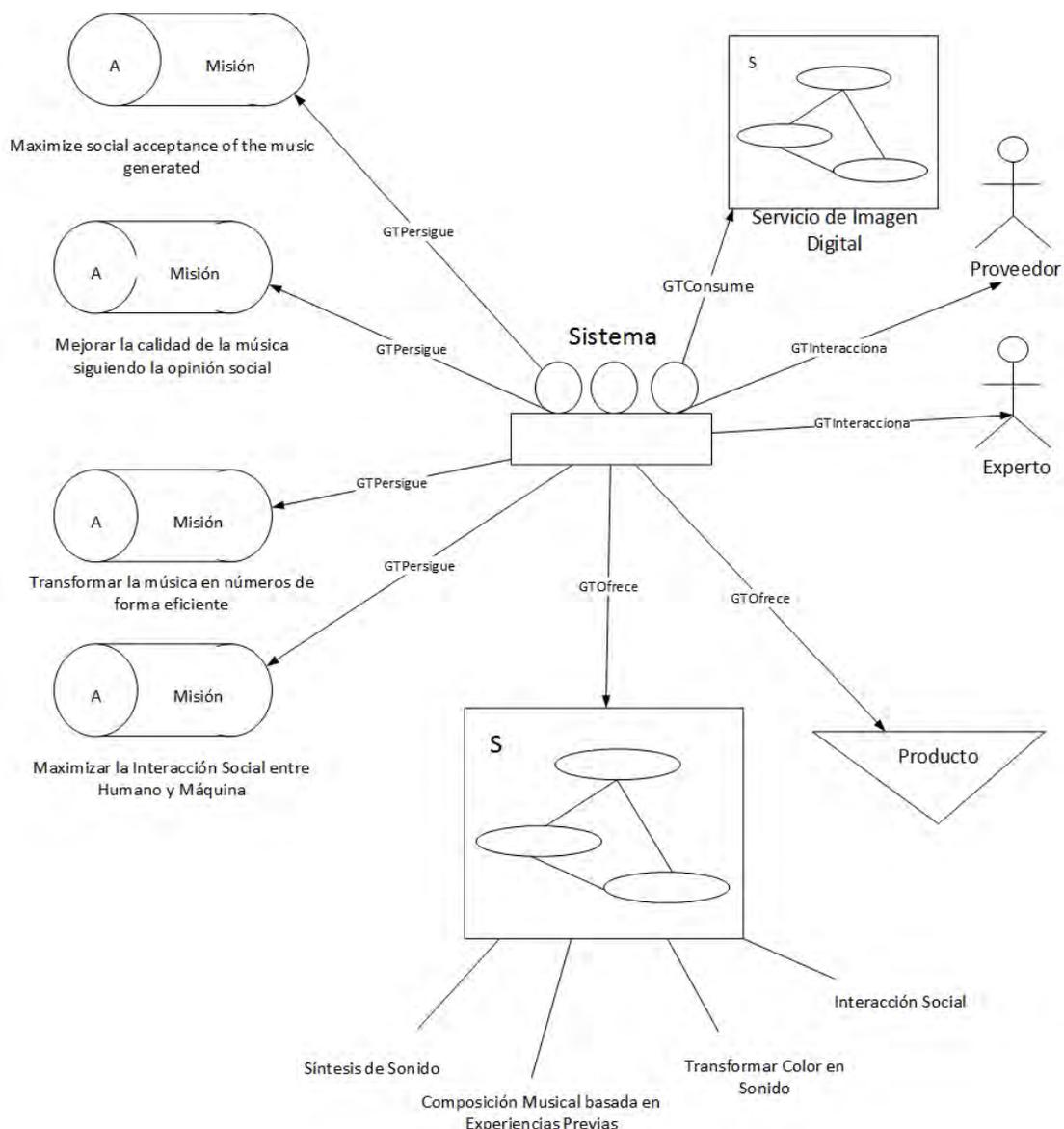


FIGURA 6.2: Representación GORMAS para la visión funcional. Esta imagen muestra las misiones de todo el sistema.

6b). El Planificador presenta los resultados musicales a los expertos (Paso 7) y recoge sus puntuaciones (Paso 8). El Planificador es responsable de considerar si las evaluaciones de estos expertos podrían almacenarse para ser utilizadas en el futuro (Paso 9).

Tenga en cuenta que el rol Planificador se implementará idealmente con una arquitectura CBR, la cual se describirá detalladamente en la siguiente sección para comprender mejor el comportamiento de dicho agente.

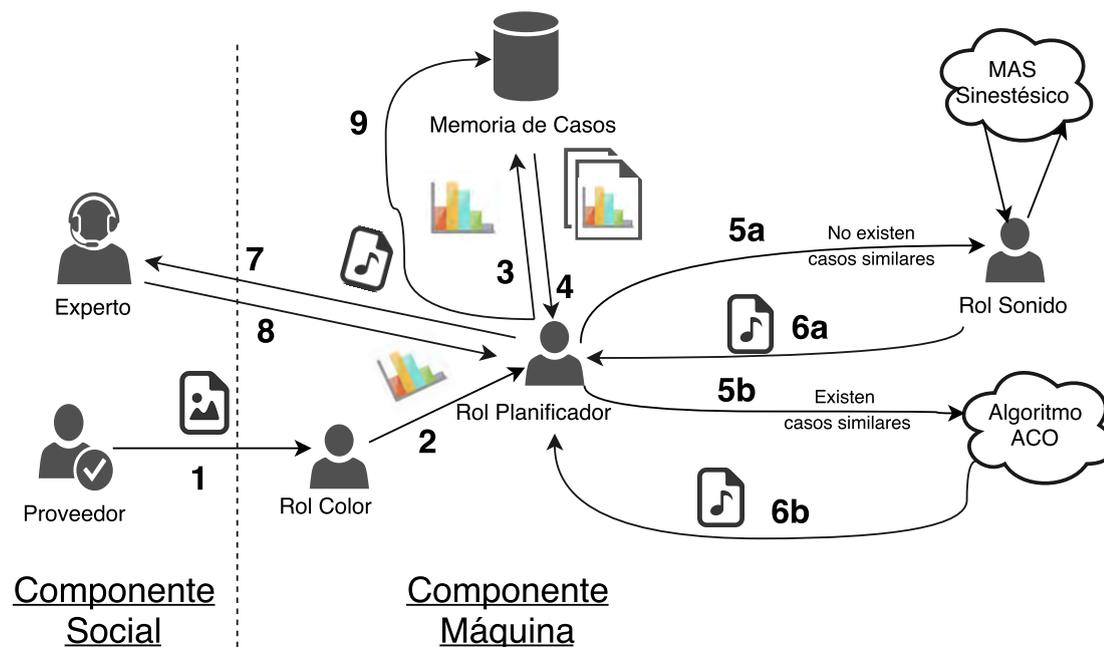


FIGURA 6.3: Flujo de trabajo del sistema basado en computación social teniendo en cuenta la organización de agentes.

### 6.3 Proceso de Aprendizaje Basado en Casos

Una de las partes esenciales del sistema aquí descrito es poder aprender del contexto. Para ello, se aplicó una arquitectura CBR a un agente perteneciente a la VO. En esta sección se hace una descripción técnica del ciclo CBR, que forma parte del rol planificador en la VO, implementado por el Agente Planificador.

Como describimos en el Capítulo 2, cualquier agente construida con una arquitectura CBR sigue las cuatro etapas del ciclo CBR, a saber, Recuperar, Reutilizar, Revisar y Retener [Bajo et al., 2007]. En nuestro caso, esta entidad está representada por un rol, que a su vez es codificado por un agente. Llamamos a este agente un agente Planificador, ya que es capaz de planificar acciones de acuerdo con la información recuperada del contexto. En particular, el agente puede seleccionar las soluciones almacenadas en su memoria que son similares a la descripción del presente caso (Paso de Recuperación). Posteriormente, adapta las soluciones anteriores para lograr una nueva solución (Paso de Reutilización). Este proceso de adaptación utiliza un algoritmo bio-inspirado para extraer una nueva solución, que debe ser evaluada en términos de calidad musical por expertos humanos (Etapa de Revisión). Finalmente, el agente almacena el caso si es útil para el futuro, basado en las puntuaciones de los expertos (Etapa de Retención). La trayectoria seguida se representa en la Figura 6.4.

Todo el proceso se basa en el concepto de caso, que puede definirse como una 3-tupla  $C = \langle P, S, R \rangle$ .  $P$  es el dominio del problema y representa el histograma de color de la

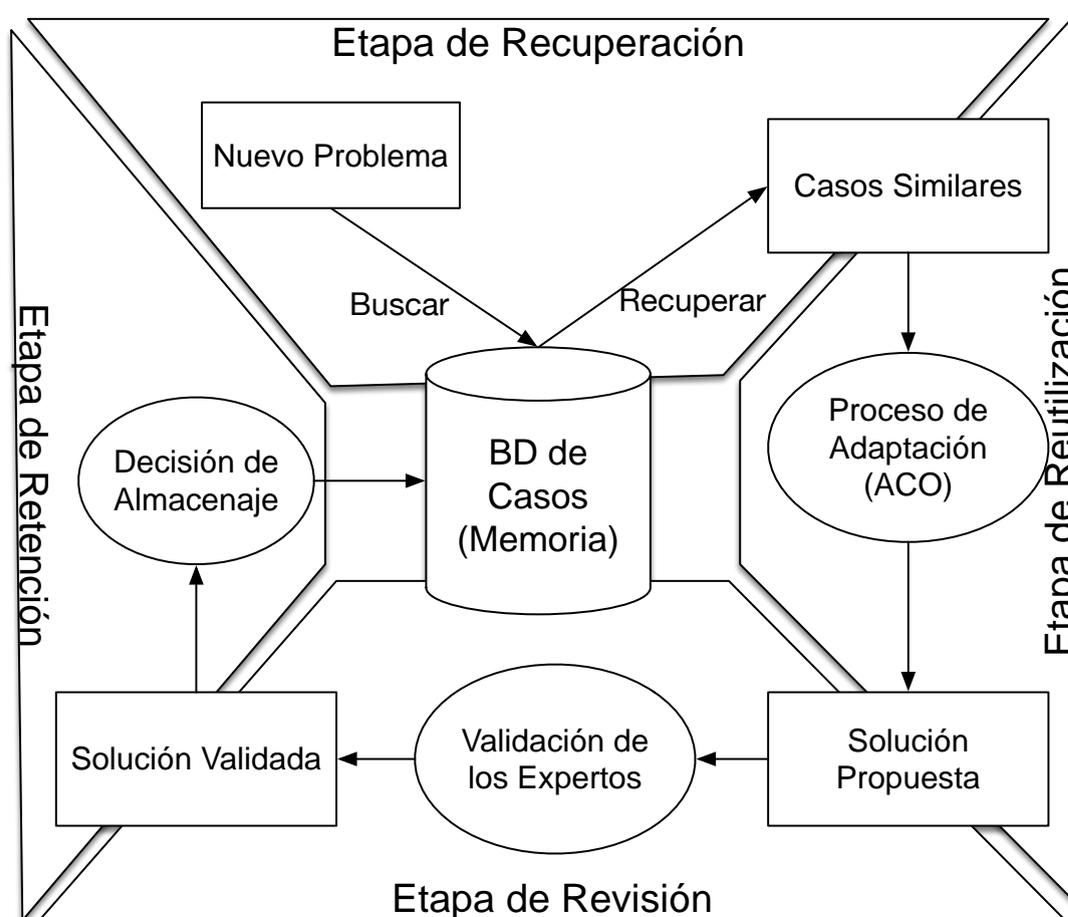


FIGURA 6.4: Esquema con el proceso CBR llevado a cabo por el Rol Planificador.

imagen.  $S$  es un conjunto de pares  $N = \langle H, D \rangle$  donde  $H$  es la nota musical considerada y  $D$  su duración rítmica. Finalmente,  $R$  representa las valoraciones subjetivas de los expertos sobre la melodía musical obtenida.

Las subsecciones siguientes detallan las técnicas aplicadas en el ciclo CBR.

### 6.3.1 Midiendo la Similitud en los Colores

En la fase de recuperación, se extraen todos los casos anteriores, es decir, todas las 3-tuplas  $C = \langle P, S, R \rangle$  almacenadas de experiencias anteriores, y el componente  $P$  se compara con la descripción de entrada de la imagen digital proporcionada por los usuarios. Para poder manejar el conjunto de colores de una imagen y poder compararlos de una forma estándar y sencilla, se decidió mejorar al Agente Color, proveyéndole de la capacidad de guardar, recuperar y analizar histogramas de color. Nuestro histograma de color se basa en el estándar visual MPEG-7 [Sikora, 2001]. Esta codificación contiene siete descriptores que pueden usarse para medir la similitud basándose en criterios

visuales. En particular, los descriptores proporcionan información sobre características elementales tales como color, textura y forma [Chang et al., 2001]. Los descriptores de color incluyen un histograma de color escalable, los colores dominantes (DCD), la estructura de color, el diseño de color y el color GoF/GoP.

Algunos estudios [Mojsilovic et al., 2002] demuestran que la percepción visual humana identifica automáticamente los colores dominantes en las regiones de la imagen, eliminando los colores de áreas muy pequeñas. Así, a nivel macroscópico, los seres humanos perciben las imágenes como una combinación de colores dominantes de diferentes regiones [Yang et al., 2008].

En MPEG-7, DCD proporciona una representación de color dominante eficaz, que describe la distribución de color en una imagen. Este descriptor de características contiene colores representativos en la codificación HSL (Tono, Saturación y Luminosidad) y el porcentaje de aparición para cada color [Yang et al., 2008].

Hay un montón de medidas para analizar la similitud entre los histogramas de color. La intersección del histograma [Chaudhary and Pithadia, 2014] mide la similitud de dos distribuciones de probabilidad (histogramas) que se superponen entre sí, con posibles valores de la intersección entre 0 (sin superposición) y 1 (distribuciones idénticas). También hay parámetros estadísticos comúnmente utilizados, como la distancia Minkowski, Chi-Cuadrado o la correlación entre histogramas [Rubner et al., 2000]. Uno de los problemas de las medidas anteriores es que necesitan que los histogramas tengan el mismo número de colores, lo que no siempre es posible para todas las imágenes [Rubner et al., 2000]. Por lo tanto, y siguiendo la recomendación en Rubner et al. [2000] hemos seleccionado la Distancia “Earth Move” (EMD) para comparar histogramas de color [Anderson et al., 2013].

Siguiendo este enfoque,  $P$  es una distribución de  $m$  colores  $P = \{(p_1, w_{p_1}), \dots, (p_m, w_{p_m})\}$ , donde  $p_i$  es el bin de color y  $w_{p_i}$  es el peso o el grado de apariencia en la imagen. De forma similar,  $Q$  contiene una distribución de color  $Q$  con  $n$  bins  $Q = \{(q_1, w_{q_1}), \dots, (q_n, w_{q_n})\}$  y  $D = [d_{p_i, p_j}]$  es la distancia entre  $p_i$  y  $q_j$ . La distancia EMD entre las distribuciones  $P$  y  $Q$  puede definirse como:

$$EMD(P, Q) = \frac{\sum_{i=1}^m \sum_{j=1}^n f_{i,j} d_{p_i, q_j}}{\sum_{i=1}^m \sum_{j=1}^n f_{i,j}} \quad (6.1)$$

donde  $f_{i,j}$  se define como el flujo entre  $P$  y  $Q$ . El flujo óptimo  $F$  se calcula como un problema de optimización lineal con las siguientes restricciones:

$$f_{i,j} \geq 0, 1 \leq i \leq m, \quad 1 \leq j \leq n \quad (6.2)$$

$$\sum_{j=1}^n f_{i,j} \leq w_{pi}, \quad 1 \leq i \leq m \quad (6.3)$$

$$\sum_{j=1}^n f_{i,j} \leq w_{pi}, \quad 1 \leq i \leq m \quad (6.4)$$

$$\sum_{i=1}^m f_{i,j} \leq w_{qj}, \quad 1 \leq j \leq n \quad (6.5)$$

$$\sum_{i=1}^m f_{i,j} \leq w_{qj}, \quad 1 \leq j \leq n \quad (6.6)$$

En la práctica, EMD proporciona el coste mínimo para transformar la distribución del histograma  $P$  en  $Q$ . Cabe señalar que los valores bajos de EMD indican imágenes muy similares. Para seleccionar esas imágenes muy similares, se calculó empíricamente un umbral.

### 6.3.2 Integrando la Inteligencia de Hormigas en el Proceso de Aprendizaje CBR

En la etapa anterior, el rol Planificador logró el primer objetivo de su arquitectura: proporcionar casos similares a la experiencia actual que está analizando. En este punto, puede surgir una situación en la que no se recuperan casos similares. En tales casos, el Planificador da el control al rol de Composición, que se encarga de generar un caso completamente nuevo sin considerar las experiencias anteriores. Los detalles completos de estos roles se proporcionan en el capítulo anterior. De lo contrario, se considera el dominio de solución  $S$  de cada caso recuperado para construir una nueva solución. En ambos casos, el resultado se revisa en el siguiente paso (Sección 6.3.3).

La solución  $S$  contiene el conjunto ordenado de notas  $N$  que representa la composición musical y la duración  $D$  de cada nota en el contexto. Cada nota puede considerarse como un plan  $p_a(t)$  de nuestro rol Planificador de acuerdo con:

$$p_a = \langle p_a(1), p_a(2), \dots, p_a(t) \rangle = \langle \langle n_1, f_1 \rangle, \langle n_2, f_2 \rangle, \dots, \langle n_t, f_t \rangle \rangle \quad (6.7)$$

El Planificador tiene como objetivo encontrar una ruta eficiente de los planes donde cada plan representa una nota individual que forma parte de la composición musical.

Supongamos que  $Pl$  son los planes de los casos originales recuperados y  $Pl'$  representan los planes generados por el agente y  $Re$  los recursos consumidos. La función de eficiencia se puede formalizar de la siguiente manera:

$$E_f = \frac{Pl \cap Pl'}{Re} \quad (6.8)$$

La ecuación fuerza la reutilización de algunos planes de las experiencias anteriores almacenadas para generar la nueva solución.  $R$  representa el conjunto de valores de aptitud que a su vez son parte del coste de usar una u otra nota. Este coste depende de la nota individual y de las tarifas globales dadas por los expertos. Consideramos cada nota como un plan para que el conjunto de planes pueda ser visto como un camino hacia un objetivo. Parece entonces lógico buscar algoritmos que optimicen las rutas entre los diferentes nodos. Entre los diferentes algoritmos, se seleccionó el algoritmo de colonia de hormigas (ACO), debido a nuestra experiencia previa en algoritmos bio-inspirados.

Tal como vimos en el Capítulo 2, las hormigas artificiales obtienen soluciones óptimas creando caminos diferentes en el espacio y representando todas las posibles soluciones  $S$ . Ellas registran sus posiciones y la calidad de sus soluciones de acuerdo con una función de fitness o coste  $f$  que debe ser optimizada. Por lo tanto, y teniendo en cuenta la representación, las hormigas artificiales construyen soluciones teniendo en cuenta la construcción gráfica  $G = (N, L)$ , donde  $H$  es el conjunto de estados posibles y los nodos en el gráfico y  $L$  son los enlaces entre los nodos  $N$ . En nuestro caso particular, cada nodo  $n_i$  es una nota musical, mientras que cada conexión es el costo de pasar de la nota  $n_i$  a la nota  $n_{i+1}$ . La función de coste es una combinación lineal de la calidad de transición  $\kappa(n_i, n_j, n_l)$  entre las notas  $n_i$  y  $n_j$ , y la diferencia de las puntuaciones globales de composiciones musicales previas  $u(n_i, n_j)$ , de acuerdo con la ecuación 6.9.

$$f(s, t) = \sum_{i=1, j=1}^N f(l_{ij}, t) \sum_{i=1, j=1}^N \Delta \cdot u(n_i, n_j) + (1 - \Delta) \cdot \kappa(n_i, n_j, n_l) \quad (6.9)$$

$l_{ij} \in L$  es el conector que une los nodos  $n_i$  y  $n_j$  y  $N$  es el número de nodos existentes en  $s$ .  $\Delta$  es una constante que equilibra la importancia entre  $\kappa$  y  $u$ , y se ha establecido empíricamente en 0.3. La calidad individual  $\kappa$  ya se ha definido y aplicado en el capítulo anterior.

Dado el conjunto de casos anteriores con notas de melodía diferentes, construimos el gráfico  $G$  considerando que cada nota en un momento específico  $i$  de un caso  $c_m$  puede ir a la nota  $n_{i+1}$  de cualquier caso  $c \in C$ . Ilustramos esta idea en la Figura 6.5.

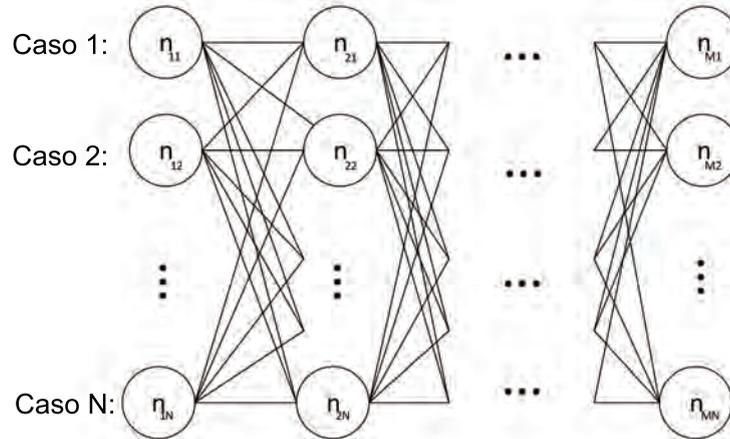


FIGURA 6.5: Ilustración de la construcción del gráfico  $G$  que representa los diferentes caminos que pueden tomar las hormigas.

Las hormigas se inician en un nodo aleatorio  $n_1$  de casos diferentes. A partir de aquí, pueden pasar de la nota  $n_i$  a la nota  $n_j$  en un momento específico  $t$  según la probabilidad  $pr$  dada:

$$pr_{ij}^k(t) = \frac{\tau_{ij}^A \cdot \sigma_{ij}^B}{\sum_{l \in N_i^k} \tau_{il}^B \cdot \sigma_{il}^B} \quad (6.10)$$

$A$  y  $B$  son constantes empíricamente establecidas, mientras que  $N_i^k$  contiene todos los nodos factibles que se pueden alcanzar desde el nodo  $n_i$ . Una hormiga completa la solución cuando llega a un nodo final, lo que significa que la secuencia contiene  $n$  nodos. Las trayectorias de feromonas necesitan ser actualizadas por una hormiga  $k$  considerando el factor de evaporación de feromonas  $\omega$  entre 0 y 1.  $\omega$  hace posible evolucionar y abandonar viejos caminos no estimulados recientemente (malas soluciones), de forma que el nivel de feromonas asociado a caminos no elegidos por las hormigas disminuya con el número de iteraciones. La actualización de los senderos de feromonas se formaliza en la Ecuación 6.11.

$$\tau_{ij}(t+1) = (1 - \omega) \cdot \tau_{ij}(t) + \sum_{k=1}^M \Delta\tau_{ij}^k(t) \quad (6.11)$$

donde  $M$  corresponde con el número total de hormigas y  $\Delta\tau_{ij}^k(t)$  es el nivel de feromona depositado en la transición  $l_{ij}$ . En este caso,  $\Delta\tau_{ij}^k(t)$  se define como:

$$\Delta \tau_{ij}^k(t) = \begin{cases} \frac{1}{f(l_{ij,t})} & : \text{if } l_{ij} \in L \\ 0 & : \text{otherwise} \end{cases} \quad (6.12)$$

donde  $L$  es el conjunto de conectores o enlaces que representan la solución  $s$  de la hormiga. Por lo tanto, la optimización de la función  $f$  resulta en un aumento de la trayectoria de feromonas en la ruta que está explorando la hormiga  $k$ . El proceso converge cuando se obtiene una solución óptima y esto se traduce en una nueva melodía basada en los casos anteriores almacenados en la memoria del Planificador. Entonces el Agente Reproductor sintetiza y reproduce este conjunto de notas para la evaluación de los expertos, que se explica detalladamente en la siguiente sección.

### 6.3.3 Recogiendo la Opinión Social

Una de las partes esenciales de este sistema es el proceso de revisión, en el cual los resultados son validados y se consideran almacenados en la memoria del sistema. La evaluación puede seguir múltiples métodos automáticos. Sin embargo, siguiendo la naturaleza de la computación social, seleccionamos la validación experta para calificar las composiciones creadas. Es importante señalar que tales evaluaciones musicales generalmente dependen de preferencias personales, formación musical o cultura social. Para extraer conclusiones válidas, se tuvo en cuenta una amplia muestra de opiniones de expertos.

Todas las opiniones de expertos tienen el mismo peso en la tasa global de cada composición, que es la media entre todas las puntuaciones en un momento dado. De ese modo, suponemos que todos los evaluadores tienen el mismo nivel de experiencia y que están dando puntuaciones honestas de acuerdo a sus preferencias. Un examen de las diferentes funciones que un experto puede tomar se deja para un estudio futuro.

El resultado musical obtenido se califica con una escala cualitativa. Esta escala contiene cinco valores categóricos, a saber, “Muy malo”, “Malo”, “Regular”, “Bueno” y “Muy bueno”. El resultado global obtenido permite decidir si la calidad del caso es lo suficientemente fuerte para su uso futuro.

### 6.3.4 Almacenamiento del Caso

A todas las composiciones musicales se les asigna una puntuación global. Nuestro rol Planificador es el responsable de decidir si se debe considerar el caso para un uso futuro. Para ello, establece dinámicamente un umbral para separar ambos tipos de

puntuaciones. Este umbral depende de los casos almacenados previamente. Si estamos almacenando casos muy malos, un caso con una puntuación global “Regular” puede ser muy interesante. Por otro lado, con casos muy buenos, un caso con puntuaciones más ajustadas no es tan interesante. Por lo tanto, el umbral se establece calculando la media para todas las tasas globales de la composición. Sin embargo, si no hay suficientes composiciones en la base de datos, el umbral se establecerá en “Regular”.

Como detallamos en una sección anterior, el sistema es dinámico y los expertos pueden dar sus valoraciones en diferentes momentos. Cada nueva puntuación puede tener un impacto positivo o negativo en la media global. Eso significa que todos los casos deben ser almacenados en nuestra base de datos incluso si su calidad no es lo suficientemente alta, ya que pueden ser válidos en un futuro. El rol Planificador actualizará la tasa global cada vez que se reciba una opinión de un nuevo experto.

## 6.4 Conclusiones

La computación social ha sido aplicada para crear interacción entre expertos y nuevos productos basados en opiniones sociales. En nuestra propuesta actual, el componente social es esencial en el proceso de producción para transformar la imagen en música, ya que motiva el inicio del proceso de cargar una obra de arte pictórica y además la opinión social se tiene en cuenta para mejorar los resultados futuros.

El sistema, implementado con una VO, extrae el histograma de color y lo utiliza para seleccionar experiencias anteriores con características similares. Estas experiencias se llaman casos, y constituyen la memoria del sistema. Los casos recuperados se adaptan utilizando un algoritmo ACO para seleccionar aquellos caminos que unen las mejores notas musicales y crear un nuevo caso basado en los anteriores. El proceso de selección considera la opinión social de los expertos musicales, así como la calidad de la música asociada a cada nota individual. El resultado final es sintetizado para la comunidad social, que se encarga de evaluar los resultados según criterios musicales y percepción de similitud entre música y colores. Estas valoraciones se utilizan para decidir si el resultado musical se puede utilizar en el futuro. El proceso principal para recuperar casos similares, adaptarlos y decidir si se pueden almacenar en la memoria es llevado a cabo por una entidad específica, el Planificador, que se construye siguiendo una arquitectura CBR.

En el siguiente capítulo se describirá cómo integrar este sistema generador de melodías junto con el de armonías, y cómo se han evaluado para verificar su calidad y eficacia.



# Capítulo 7

---

## Caso de Estudio: Un Framework Compositor

---



VNiVERSiDAD  
DSALAMANCA



# Caso de Estudio: Un Framework Compositor

---

## 7.1 Introducción

El objetivo principal de este trabajo es construir una sociedad humano-agente a través de un framework que integre diferentes mecanismos, sistemas, arquitecturas y modelos para que agentes virtuales y humanos colaboren y generen música. A partir de una motivación sugerida por el usuario, y que puede consistir en una imagen o una serie de parámetros de preferencia, el framework es capaz de generar una composición musical que se ajuste a estos datos de entrada. Por tanto, los resultados finales dependerán del contexto en el que han sido generados, entendiendo como contexto los parámetros de entrada, el entrenamiento del sistema y la interacción con el usuario.

El framework integra dos componentes principales que pueden interaccionar con el usuario. El primer componente está descrito en el Capítulo 4, y consiste en un módulo que es capaz de generar progresiones de acordes tonales de acuerdo a los criterios de entrada que el usuario ha decidido establecer. El segundo componente es capaz de crear composiciones musicales melódicas que siguen unos cánones tonales, pero que están basadas en un parámetro de entrada inusual, como es una imagen digital.

Uno de los puntos a explotar es que este sistema ha sido diseñado para funcionar de forma social, y funciona por tanto imitando una sociedad humano-agente. Esto significa que los usuarios pueden interaccionar en el propio framework, valorando no solo sus propias composiciones, sino melodías que otros usuarios han generado. Además, aunque los resultados han sido generados utilizando técnicas basadas en la inteligencia artificial, siempre es necesario un componente humano para validar los resultados y mejorar el sistema. De este modo, la interacción humano-agente se convierte en punto esencial para el éxito de nuestro framework.

La propuesta pretende utilizar la motivación que el ser humano proporcione al sistema para iniciar un proceso creativo y transformarla en un resultado musical de acuerdo a las preferencias de los distintos usuarios que conforman la sociedad. Es importante destacar que la generación de contenido creativo como la música es un problema desafiante debido al dinamismo del contexto, ya que los usuarios pueden cambiar su motivación en cualquier momento y esperan resultados valiosos según sus criterios, lo cual es también muy subjetivo y puede depender de factores como su cultura o su estado de ánimo.

Para comprender la complejidad de este trabajo, en capítulos anteriores se ha descrito en profundidad cada parte de nuestro framework. En este capítulo, describiremos cómo se han integrado los diferentes componentes y cómo se favorece la interacción entre humano y agente para generar y mejorar los resultados musicales. De la misma forma, se ha llevado a cabo una evaluación de cada una de las partes de forma individual, para analizar la calidad de los resultados en cada uno de los componentes. Finalmente, se ha realizado un estudio del framework total para estudiar la tasa de éxito o de fracaso a la hora de interactuar con el sistema global.

## 7.2 Funcionamiento del Framework

En los capítulos anteriores, se han detallado tres subsistemas interdependientes que son capaces de generar resultados musicales por sí solos, a saber: generador de armonías, generador de melodías y componente de aprendizaje. A pesar de la integración de diferentes estrategias evolutivas que les dotan de una pseudo-inteligencia, estas tres partes necesitan de la interacción humana para generar resultados musicales de calidad.

Como hemos visto en anteriores capítulos, para diseñar una VO hemos seguido la metodología GORMAS, que propone cuatro dimensiones (estructural, funcional, ambiental y normativa) para describir las organizaciones virtuales [Argente et al., 2009]. Cada dimensión define un procedimiento cuyo resultado es un esquema global o artefacto. GORMAS analiza profundamente dentro de la organización virtual. Aquí hemos representado los artefactos más relevantes para definir los roles y las tareas.

Los roles aquí representados engloban no solo los identificados en el Capítulo 6 (Proveedor, Color, Planificador, Compositor, Reproductor y Experto), sino también los relativos al framework global (Interfaz y Rol Progresión).

Los roles descritos pueden ser implementados por diferentes agentes, y estos agentes pueden incluso cambiar con el tiempo. El esquema de los roles del sistema se presenta en la Figura 7.1. Cada agente juega un papel, que se representa en la figura con la flecha “WFPlays”. Tanto los agentes como los roles son elementos de la organización general

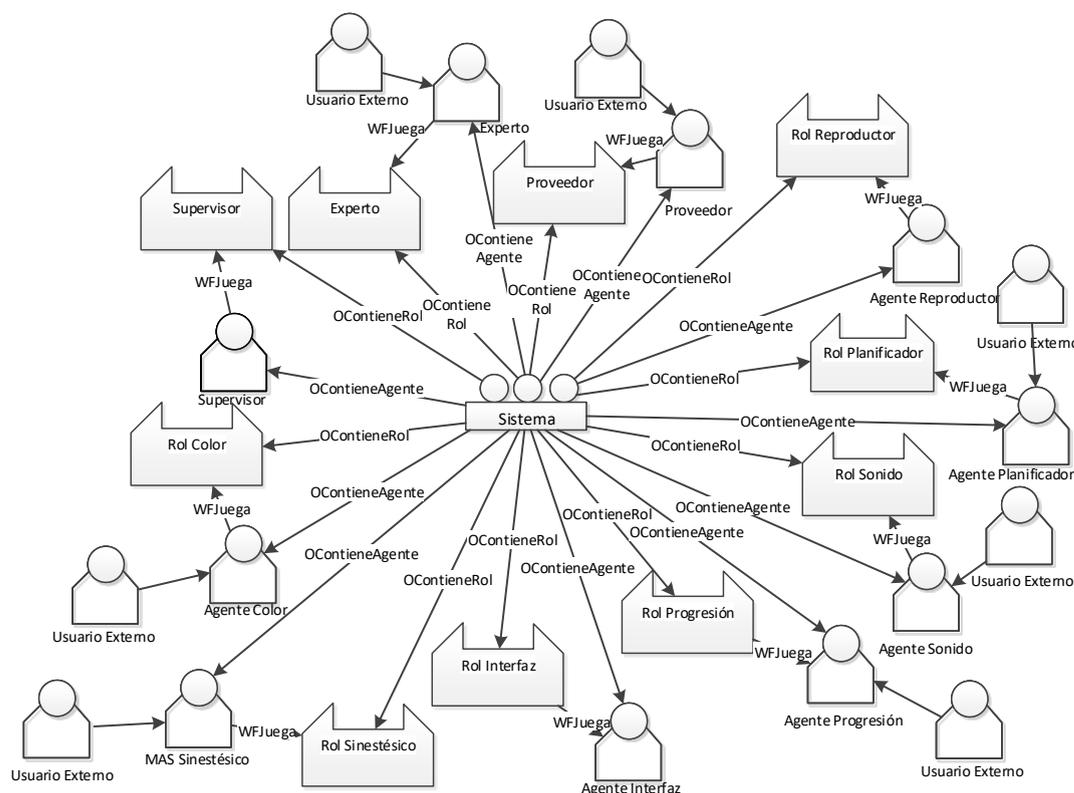


FIGURA 7.1: Roles del sistema en una Organización Virtual

(representada por la forma central). Por lo tanto, están vinculados al sistema por las flechas “ OContainsRole ” y “ OContainsAgent ”. Es esencial considerar a algunos agentes como agentes externos, es decir, entidades que pueden ser implementadas por varios elementos y que pueden cambiar con el tiempo. Un agente externo puede ser un ser humano que interactúa con el sistema o una entidad de software inteligente que implementa un modelo de razonamiento. Además, es necesario tener en cuenta que los agentes externos pueden desempeñar diferentes funciones. En particular, tenemos seis funciones desarrolladas por agentes externos, a saber, el rol de sonido, el rol sinestésico, el rol planificador, el rol de proveedor y de experto y el rol progresión. Este diseño de agentes abiertos da flexibilidad al sistema para implementar, mejorar, agregar o incluso reemplazar alguna funcionalidad. También es importante destacar la misión de los roles de proveedor y experto como usuarios humanos de nuestro framework.

El siguiente paso en la metodología GORMAS consiste en identificar las metas del framework, los recursos necesarios para llegar a ellas y los servicios que el framework puede proporcionar al entorno externo. En este caso, se consideran seis misiones: maximizar la calidad de los acordes generados en cada iteración, maximizar la aceptación de la progresión armónica, maximizar la exactitud para extraer el histograma de color, maximizar la aceptación social de la melodía generada, maximizar la eficiencia para

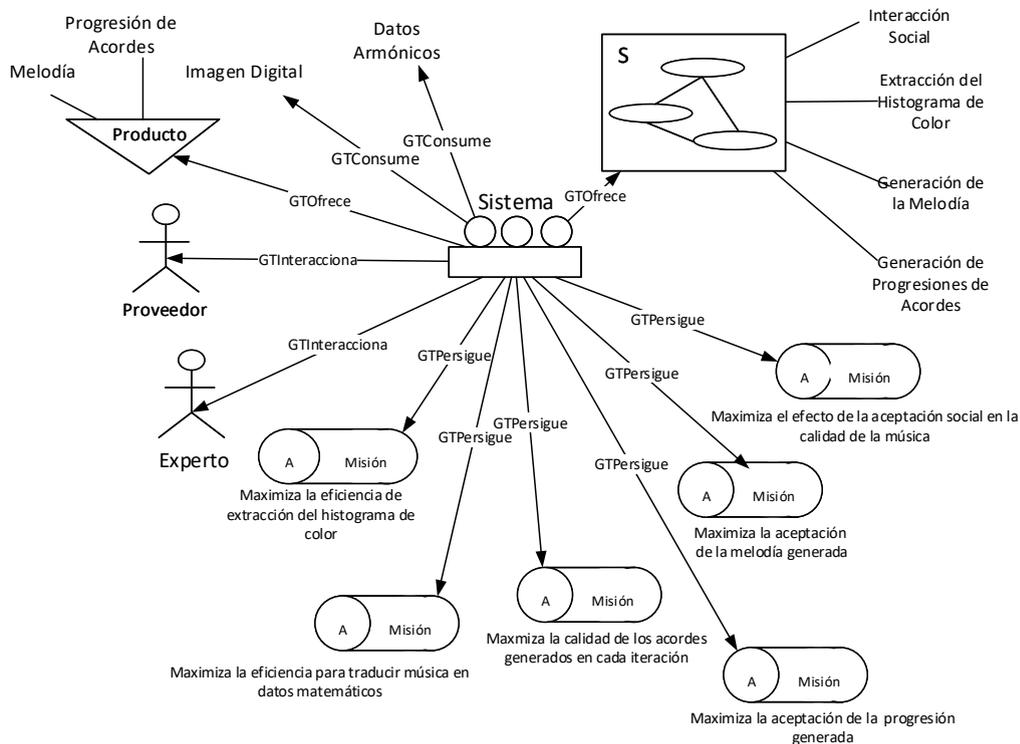


FIGURA 7.2: Representación GORMAS para la visión funcional. Esta imagen muestra las misiones de todo el sistema.

traducir la información numérica a música y maximizar el efecto de las interacciones sociales en el framework en general. El recurso principal para lograr las misiones es la imagen digital cargada por los proveedores. Además, el sistema puede proporcionar los siguientes servicios: generación de progresión de acordes, creación de melodía, extracción de histograma de color e interacción social. Esta vista de funcionalidad descrita también se representa en la Figura 7.2.

Este diseño basado en organizaciones de agentes nos permite obtener un framework flexible y modular, donde algunos agentes pueden ser modificados o incluso sustituidos de forma sencilla, sin necesidad de modificar el resto del comportamiento de las organizaciones.

Tal como hemos visto con los esquemas GORMAS, para colaborar de forma eficiente con el usuario y generar así una sociedad humano-agente, se ha integrado cada componente dentro de una plataforma que soporta el desarrollo de diversas arquitecturas y metodologías, como son las organizaciones virtuales o los algoritmos bio-inspirados. La Figura 7.3 ilustra cómo interactúan los módulos con los usuarios del sistema. Algunos agentes con un rol secundario en nuestro sistema, han sido omitidos para mayor claridad.

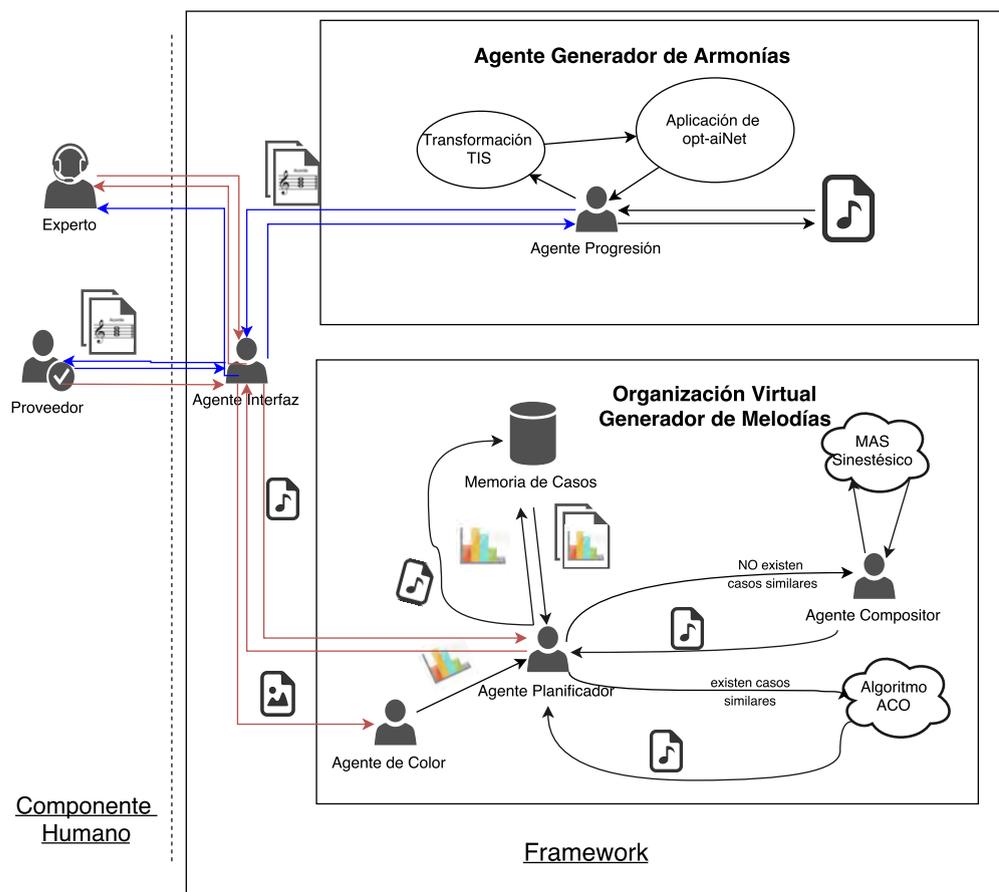


FIGURA 7.3: Representación general del framework. Las flechas azules indican la interacción de los usuarios con la parte generadora de armonías. Las flechas rojas indican interacciones de usuarios con la parte generadora de melodías.

Inicialmente, el usuario accede a nuestro framework para elegir entre componer melodías o componer armonías a través del agente Interfaz, que recoge toda la comunicación entre el usuario y el resto del framework. En el primer caso, el agente Interfaz pasa el control al sistema generador de armonías, que está implementado en un agente llamado Agente Progresión, encargado de crear una progresión de acordes a partir de una tonalidad y un acorde de entrada que el usuario debe introducir. Asimismo, cada vez que el Agente Progresión realiza una iteración, el sistema inmunológico desarrollado devuelve un máximo de cuatro posibles acordes que pueden incorporarse a nuestra progresión. La elección de uno u otro solamente dependerá del usuario, quien debe seleccionar qué opción le parece más adecuada de acuerdo a sus preferencias, representado en la figura mediante flechas azules. Una vez que el acorde se incorpora, el agente reactiva el sistema inmunológico para buscar nuevos candidatos que son posibles acordes a añadir a la progresión. La progresión final puede ser valorada por los usuarios expertos a través del agente Interfaz.

En el segundo caso, el agente Interfaz cede el control a la organización virtual

desarrollada en el Capítulo 5 y perfeccionada en el Capítulo 6, responsable de generar el resultado melódico a partir de una imagen digital que se le pide al usuario como parámetro de entrada. Una vez generado, la organización virtual muestra el resultado al agente Interfaz, y en consecuencia al usuario, quien debe evaluarlo de acuerdo a sus preferencias personales. Esta evaluación es incorporada a la memoria de la VO para su posterior uso en futuras composiciones. Todas las interacciones con los usuarios están señaladas, en este caso, con las flechas rojas.

Es necesario destacar que todas las composiciones están disponibles para que otros usuarios las puedan valorar de acuerdo a criterios musicales o a sus propias preferencias. No obstante, estas evaluaciones solo servirán para la mejora del sistema en el caso del sistema generador de melodías.

Para validar el framework total desde un punto de vista computacional, se ha utilizado PANGEA [Zato et al., 2013] como herramienta para implementar un prototipo funcional. PANGEA es una plataforma orientada a servicios que permite la implementación de una Organización Virtual y se aprovecha al máximo de la distribución de recursos mediante la implementación de servicios estándar como Servicios Web. Esta plataforma para la construcción automática de organizaciones de agentes inteligentes está orientada hacia el servicio, lo que significa que puede integrar diferentes herramientas modeladas con agentes que consumen servicios web, independientemente de su ubicación física o implementación [Zato et al., 2013].

La plataforma acepta la programación Java e incluye un protocolo de comunicación útil para nuestra propuesta, una herramienta de depuración y flexibilidad total para implementar diferentes modelos de agentes, topologías y organizaciones [Zato et al., 2013].

Una vez diseñado e implementado, es necesario demostrar la validez del framework. La validación de este framework puede tener un componente computacional (evaluando la eficiencia computacional del sistema) o creativo, si lo que queremos es validar los resultados musicales. La evaluación en este último caso sigue siendo un reto en la generación de música con sistemas expertos debido a la subjetividad inherente de los oyentes humanos [Pearce et al., 2002]. La mayoría de las investigaciones utilizan medidas objetivas y algún método de optimización sin evaluación humana en absoluto, mientras que unos pocos utilizan un experto humano para evaluar los resultados. López-Ortega and López-Popa [2012] discuten la calidad del sistema aplicando teóricamente algunos conceptos creativos, como la deliberación y la espontaneidad. Herremans and Sörensen [2013] presentan un experimento estadístico exhaustivo para validar la eficiencia de acuerdo con las reglas diseñadas en el sistema. Herremans et al. [2015] comparan medidas objetivas con un experto.

Hay ejemplos de evaluación subjetiva por un grupo de oyentes humanos que no son expertos. Delgado et al. [2009] realizar una prueba de escucha preliminar para evaluar su aplicación. Pearce and Wiggins [2007] proponen una prueba de escucha y un estudio con medidas estadísticas. Collins et al. [2016] también realizan una prueba de escucha y un estudio comparativo entre el sistema actual y un trabajo anterior. En este capítulo se recurrirá a varios tests de escucha para evaluar los resultados musicales de cada módulo.

Por último, se evaluará el framework desde un punto de vista computacional, realizando un estudio comparativo entre varios sistemas informáticos, así como diversos análisis comparativos para destacar las ventajas de las diferentes partes de nuestro framework frente a otras propuestas.

### 7.3 Evaluación del Generador de Armonías

El objetivo de la evaluación del generador de armonías es triple: validar primero los vectores TIV como representación, validar la función objetivo que diseñamos y finalmente validar el uso de un sistema inmunológico (AIS) en el sistema global. La validación de los vectores TIV se centrará en las propiedades topológicas del espacio TIS independientemente de la tonalidad. Así, se seleccionaron cuatro escalas diferentes en los modos mayor y menor para evaluar si la calidad de las progresiones de acordes generadas dependen de la tonalidad. Además, comparamos los acordes generados con otros vectores TIV y otras representaciones.

En este módulo, la generación de progresiones de acordes se formula igual que un problema de búsqueda, donde una función objetivo está diseñada para asociar un valor a cada acorde candidato dependiendo de una progresión de acordes de entrada dada. El valor objetivo de los acordes candidatos debe reflejar la calidad de este acorde para seguir la progresión actual. Para validar la función objetivo, debemos evaluar si los valores objetivos más bajos están asociados con mejores acordes candidatos que los valores objetivos más altos. Seleccionamos acordes con valores objetivos que muestran la función objetivo desde valores bajos (acordes buenos) hasta valores altos (acordes malos). Los pesos para cada término de la función objetivo en la Ecuación 4.5 se establecieron empíricamente en  $D = 6$ ,  $X = 0,1$ ,  $L = 1$ , y  $F = 1,5$ , asignando más importancia relativa a la distancia desde el acorde anterior, y menor a la disonancia.

El algoritmo Opt-aiNet es capaz de encontrar múltiples mínimos locales de la función objetivo mientras se preserva la diversidad. El acorde con el valor objetivo más bajo se consideraría el más apropiado para seguir la progresión de acorde actual. Sin embargo, otras buenas soluciones candidatas darían lugar a múltiples opciones alternativas con

buena calidad. Estas soluciones candidatas tienen diferentes valores objetivos, algunos más bajos que otros (que indican acordes potencialmente mejores). Queremos evaluar si los acordes seleccionados por opt-aiNet son considerados mejores candidatos que los acordes que fueron descartados.

Se realizó una prueba de audición para evaluar la calidad perceptual de los acordes propuestos por el AIS en el espacio TIS. Para cada tonalidad, escogimos 7 acordes candidatos seleccionados por opt-aiNet y 8 acordes adicionales que fueron rechazados, muestreando la función objetivo regularmente desde el valor objetivo más bajo al más alto.

### 7.3.1 Test de Escucha

Para validar el proceso de generación de acordes realizamos una prueba de audición, seguida de un análisis estadístico y un estudio comparativo.



FIGURA 7.4: Los acordes utilizados como secuencia de entrada para generar el siguiente candidato.

La prueba de audición presentó una secuencia de dos acordes de entrada seguida de un tercer acorde generado por el AIS y pidió al oyente que evaluara el grado de adaptación y calidad del tercer acorde considerando la entrada (los dos acordes anteriores). Se incluyeron acordes seleccionados por el AIS así como acordes rechazados por él para validar la función objetivo. Se utilizaron cuatro progresiones de acordes diferentes, dos en un modo mayor (Do mayor (CM) y Sol mayor (GM)) y dos en un modo menor (Do menor (Cm) y Sol menor (Gm)), tal y como se muestran en la Figura 7.4. El primer acorde en cada progresión fue siempre la tónica en posición fundamental para estabilizar la tonalidad [Agmon, 1995] porque la tónica determina la base tonal de la música. Además, las tríadas en posición fundamental tienen un sentido más firme del centro tonal, pudiendo diferenciar los modos mayor y menor dentro de la misma tonalidad solo escuchando este acorde. El segundo acorde es siempre diferente para cada tonalidad, para representar diferentes funciones armónicas. Los acordes para la escala de Do Mayor son Do Mayor (tónica) seguido de La menor (subdominante). En la escala de Sol Mayor, el primer acorde es la tónica (Sol Mayor) seguido por Re Mayor (dominante). En Do menor se presenta el acorde de tónica seguido de Sol Mayor (dominante). Finalmente, en la tonalidad de Sol menor presentamos el acorde de tónica Sol menor seguido de Do

menor (subdominante). La prueba de audición está disponible en línea en la siguiente dirección <sup>1</sup> y consta de cuatro listas de reproducción con las progresiones de acordes presentadas aleatoriamente. Cada progresión de acordes se puede reproducir varias veces antes de evaluarlo. Se les pidió a los oyentes que evaluaran su grado de satisfacción con el tercer acorde de acuerdo a la consonancia y a su relación con el primero y el segundo acorde usando las siguientes calificaciones: muy bueno (+2), bueno (+1), aceptable (0), malo (-1) o muy malo (-2). En total, 45 personas tomaron la prueba, entre los cuales ocho declararon no tener ninguna formación musical, diecinueve se consideraban aficionados, y dieciocho músicos profesionales.

Esperamos que la función objetivo refleje la calidad perceptual de las configuraciones tonales, de manera que los acordes con valores objetivos inferiores correspondan a mejores candidatos que valores superiores, independientemente de la tonalidad. Asimismo, queremos verificar la calidad de los candidatos seleccionados por opt-aiNet y compararlos con los acordes descartados. Esperamos que los acordes seleccionados por opt-aiNet sean valorados positivamente por los oyentes, mientras que los acordes descartados sean clasificados negativamente.

### 7.3.2 Resultados y Discusión

En primer lugar, abordamos la validación de la función objetivo  $P$  como un proxy para la evaluación perceptiva de la calidad del acorde añadido a la progresión. La Figura 7.5 muestra gráficas de la media y la desviación estándar de las valoraciones subjetivas frente a los valores objetivos de los acordes incluidos en la prueba de audición. En general, los acordes con valores objetivos más bajos fueron calificados como mejores candidatos para todas las tonalidades probadas. Las puntuaciones subjetivas representadas con un triángulo sólido son acordes candidatos seleccionados por opt-aiNet, mientras que las calificaciones marcadas como puntos corresponden a acordes rechazados por el algoritmo. Todos los acordes seleccionados por opt-aiNet se encuentran por encima de la línea horizontal que corta 0 o *aceptable*, indicando que opt-aiNet selecciona los acordes que se consideran *aceptables*, *buenos* o *muy buenos* candidatos para añadir en la progresión de acordes dada. La mayoría de los acordes desechados se encuentran por debajo de la calificación de *aceptable*.

Las gráficas también muestran la regresión lineal de las calificaciones subjetivas en función de los valores objetivos con una línea recta, junto con los valores estadísticos calculados a partir de los datos, a saber, el p-valor, el coeficiente de determinación  $R^2$  y el coeficiente de correlación de Pearson  $\rho$ . El análisis estadístico sugiere que la función

<sup>1</sup><http://form.jotforme.com/form/52522142163343>

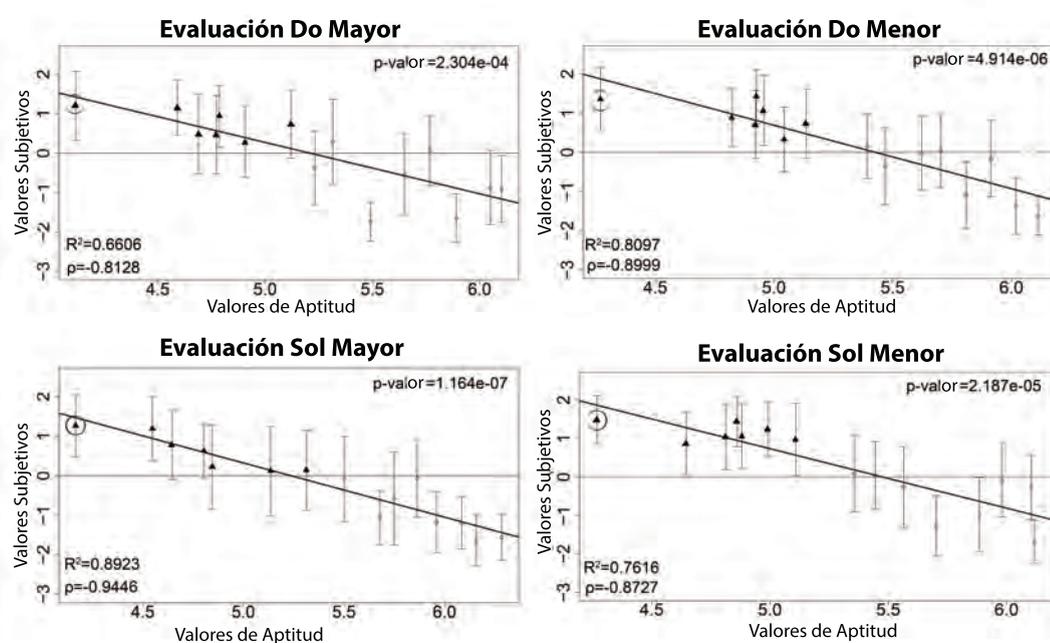


FIGURA 7.5: Gráfico de la evaluación subjetiva (puntuaciones de los oyentes) en función del valor objetivo. Cada punto representa la evaluación del tercer acorde creado por el sistema, considerando la progresión de acordes dada como entrada. El eje horizontal es el valor objetivo y el eje vertical es la media de las puntuaciones que dieron los oyentes.

La desviación estándar se muestra para cada punto usando barras verticales.

objetivo capta la calidad perceptual de los acordes. Los p-valores están todos por debajo del umbral de 1% para la hipótesis nula, los altos  $R^2$  indican que la regresión se ajusta bien a los datos y los coeficientes de correlación de Pearson  $\rho$  indican que las valoraciones subjetivas se correlacionan fuertemente con los valores objetivos.

TABLA 7.1: Comparación de las estadísticas resultantes de la organización de las valoraciones subjetivas de la Figura 7.5 en función de las medidas de distancia de otras representaciones de acordes. Véase el texto para la explicación de cada medida de distancia.

		$P(T_3)$	$d(T_2, T_3)$	$d(C_2, C_3)$	$d(D_2, D_3)$	$d(H_2, H_3)$
C Mayor	p-valor	<b>2.3e-04</b>	0.9591	0.2485	0.0111	0.7481
	$R^2$	<b>0.6606</b>	0.0002	0.1010	0.4022	0.0082
	$\rho$	<b>-0.8128</b>	0.0145	-0.3177	-0.6342	-0.0906
C Menor	p-valor	<b>4.9e-06</b>	0.1266	0.7907	0.0155	0.4005
	$R^2$	<b>0.8097</b>	0.1701	0.0056	0.3734	0.0549
	$\rho$	<b>-0.8999</b>	-0.4124	0.0749	-0.6111	0.2343
G Mayor	p-valor	<b>1.2e-07</b>	0.8374	0.4520	0.2647	0.6679
	$R^2$	<b>0.8923</b>	0.0034	0.0442	0.0946	0.0146
	$\rho$	<b>-0.9446</b>	-0.0580	-0.2102	-0.3076	-0.1209
G Menor	p-valor	<b>2.2e-05</b>	0.6377	0.1501	0.0867	0.8723
	$R^2$	<b>0.7616</b>	0.0176	0.1525	0.2090	0.0021
	$\rho$	<b>-0.8727</b>	0.1326	-0.3906	-0.4572	-0.0454

Para validar el espacio TIS como representación, utilizamos la distancia euclídea del acorde anterior como medida de calidad objetiva con diferentes representaciones de acordes. La Tabla 7.1 muestra una comparación de las estadísticas resultantes de la organización de las valoraciones subjetivas de la Figura 7.5 en función de las medidas de distancia de otras representaciones de acordes.  $P(T_3)$  es la medida objetiva de la Ecuación 4.5. Para las otras columnas,  $d(\dots)$  representa la distancia euclídea en la Ecuación 3.2b, los subíndices indican la posición del acorde y la letra representa la representación del acorde. Por ejemplo,  $d(T_2, T_3)$  es la distancia euclídea entre el vector TIV para cada segundo acorde de Fig 7.4 y cada tercer acorde utilizado en la prueba de audición.  $C$  representa el vector cromático,  $D$  representa la representación diatónica de Fig 3.1 usada en el Capítulo 3, y  $H$  representa la representación en seis dimensiones propuesto por Harte et al. [2006].

Una comparación de los análisis estadísticos en la Tabla 7.1 indica que la función objetivo  $P(T_3)$  usada para evaluar la calidad de los acordes candidatos en este trabajo se correlaciona fuertemente con las valoraciones subjetivas de la prueba de audición, mientras que la distancia euclídea entre diferentes representaciones de acordes no es estadísticamente significativa. La distancia euclidiana entre acordes no es un buen indicador de calidad de acordes para las representaciones probadas, bajo las condiciones de la prueba de audición realizada. En particular, la columna  $d(T_2, T_3)$  en la Tabla 7.1 indica que sólo la distancia euclidiana no es suficiente para elegir el siguiente acorde en el espacio TIS, justificando los otros términos en la Ecuación 4.5, tal y como predijimos al final de la Sección 4.3.1.

En Bernardes et al. [2016] se propone generar automáticamente progresiones de acordes utilizando las propiedades topológicas del espacio TIS. En su trabajo, los acordes candidatos se clasifican de forma determinista de acuerdo a una partitura que utiliza la disonancia, la distancia desde el acorde anterior y de la tonalidad. Posteriormente, a cada acorde se le asigna una probabilidad de ser agregada a la progresión actual que es proporcional a la puntuación. Un acorde se agrega a la progresión actual usando estos valores de puntuación. Cada gráfica de la Figura 7.5 resalta un acorde con un círculo. Este acorde destacado recibiría la puntuación más alta en el método desarrollado por Bernardes et al. [2016], por lo que es un candidato muy probable para ser seleccionado.

Las soluciones candidatas en el enfoque de Bernardes et al. [2016] se clasifican sólo de acuerdo con la puntuación. Los acordes con puntuaciones similares probablemente estarán muy juntos en el TIS. Consecuentemente, acordes similares serán más probables de ser agregados a la progresión actual en este caso. La propiedad de mantenimiento de la diversidad en opt-aiNet surge de la medida de afinidad entre las células de la red. La afinidad se mide como la distancia euclidiana y el algoritmo elimina individuos dentro

de un umbral de distancia  $\sigma$  a las mejores soluciones candidatas en cada iteración. En consecuencia, el algoritmo mantiene las mejores soluciones candidatas que están por lo menos  $\sigma$  separadas unas de otras. En este trabajo, este acorde destacado tiene el valor objetivo más bajo y se clasificó como el mejor para cada clave probada. Sin embargo, opt-aiNet propone otros acordes de candidatos que son diferentes del mejor candidato y todavía son clasificados positivamente por los oyentes. Opt-aiNet mantiene estos acordes candidatos, que son menos triviales de encontrar con los métodos tradicionales, debido a la característica de mantenimiento de la diversidad.

En un trabajo anterior [Navarro et al., 2015], utilizamos un algoritmo heurístico para buscar acordes candidatos, dando lugar a un espacio de búsqueda con diferentes propiedades topológicas, tales como las relaciones de cercanía entre los acordes. La codificación previa sólo es capaz de representar la escala diatónica en Do mayor. El TIS permite la generación de progresiones de acordes en cualquier tonalidad. La Figura 7.5 no muestra diferencias significativas entre las escalas utilizadas en la prueba de audición, lo que indica que la función objetivo es independiente de la tonalidad en el espacio TIS. Además, los vectores TIV también pueden representar acordes no diatónicos, a diferencia de nuestra codificación en el capítulo anterior, dando lugar a candidatos más diversos que pueden ser musicalmente interesantes desde un punto de vista armónico, como los representados en la Figura 7.6.



FIGURA 7.6: Ejemplo musical de un candidato diatónico y uno no diatónico. Ambos son seleccionados por el algoritmo y clasificados como buenos candidatos en la prueba de audición.

## 7.4 Resultados del Sistema de Generación Melódica

En este segundo módulo, el objetivo de la evaluación pretende validar la función que mide la calidad musical que diseñamos y validar que la transformación de color en sonido se ha realizado de forma exitosa. La validación de la función objetivo se basará en la evaluación de la calidad de los resultados obtenidos en cuanto a propiedades musicales. Por su parte, la transformación de color en sonido tratará de evaluar si la melodía generada puede recoger de alguna forma los colores de la imagen digital utilizada como entrada.

En este módulo, la generación de notas puede considerarse un problema de búsqueda, donde cada candidato es seleccionado de acuerdo a su calidad musical siguiendo los valores de la función de aptitud. Esta función debe capturar la calidad de la nota para seguir la línea musical actual. Para validar la función de aptitud, analizamos si valores de aptitud más bajos corresponden a melodías mejores que los valores superiores, ya que la función es de minimización. Así, seleccionamos algunas melodías con valores objetivos más bajos (melodías buenas) y otras con valores objetivos más altos (malas melodías). La función de aptitud global de una melodía se mide mediante el cálculo de la media aritmética de todos los valores de aptitud obtenidos para cada nota individual en la melodía.

La evaluación consistirá en comprobar si los valores de la función de aptitud capturan la calidad musical de los sonidos. Para ello, se diseñó un experimento preliminar con diferentes fragmentos musicales compuestos por el sistema, dejando aparte las imágenes. Este primer experimento fue presentado como un test de escucha, con fragmentos musicales en diferentes tonalidades, en particular, tres fragmentos en Do Mayor, dos en Fa# Mayor, uno en Reb Mayor, uno en Re Mayor y un extracto en Sol Mayor. Al igual que antes, la melodía comienza siempre con la nota tónica para establecer la clave porque la tónica determina la base tonal de la música [Agmon, 1995]. Las notas finales de la melodía también están relacionadas con la tónica para transmitir la sensación de final musical. La prueba de audición se puede encontrar en línea en este enlace <sup>2</sup>. Se pidió a los oyentes que evaluaran la calidad general de la melodía según los estándares tonales. Cada fragmento musical puede ser calificado de 1 (muy malo) a 5 (muy bueno) según la opinión de los oyentes (los oyentes son considerados expertos musicales con más de 6 años de formación musical). Finalmente, 44 expertos musicales realizaron esta prueba de audición y evaluaron los fragmentos musicales.

Esperamos que la función objetivo refleje la calidad perceptual de las melodías con valores objetivos inferiores correspondientes a candidatos mejores que valores superiores independientemente de la tonalidad.

Para estos datos categóricos, usamos la prueba del chi-cuadrado para demostrar que los resultados reflejan la calidad de las melodías según las valoraciones subjetivas dadas por los oyentes. Es importante notar que podemos obtener valoraciones muy subjetivas ya que todos los oyentes pueden tener diferentes interpretaciones de las piezas musicales y diferentes gustos musicales. Por lo tanto, en este caso hemos considerado el análisis de la mediana y la moda como datos estadísticos útiles en nuestro análisis.

La Tabla 7.2 recoge los valores estadísticos calculados a partir de los datos, el p-valor para el parámetro chi  $\chi^2$  la mediana  $M_e$  y la moda  $M_d$ . El análisis estadístico sugiere

<sup>2</sup><https://form.jotformeu.com/70191628235354>

que la función objetivo capta la calidad perceptual de las melodías. En particular, el valor del p-valor de  $\chi^2$  nos lleva a rechazar la hipótesis nula, que implica la existencia de una relación entre los valores objetivos y las puntuaciones subjetivas. La mayoría de los oyentes piensa que los valores son buenos (según el  $M_d$ ) y la mediana indica que al menos la mitad de las calificaciones obtenidas se califican como “Buenas” o incluso mejor. Los resultados estadísticos nos llevan a concluir que la función de aptitud o fitness capta bastante bien la calidad de la percepción musical.

El objetivo de la segunda parte del experimento es medir la correlación entre la composición de color de una imagen digital y los sonidos. Aquí, la imagen juega un papel importante porque es una fuente de inspiración que genera la música. Una nueva prueba de audición fue diseñada para calificar la percepción musical de la música y el sonido. En este caso, se presentaron los mismos fragmentos musicales junto con la imagen digital que representa el origen de la composición. Los expertos deben calificar los fragmentos musicales entre 1 (sin similitud en absoluto) y 5 (muy similar) según su percepción personal de similitud entre la música y la imagen. La prueba de audición se puede encontrar en el siguiente enlace <sup>3</sup>.

Para nuestro test, presentamos ocho entradas diferentes, como se muestra en la Figura 7.7. Para ello, se invitó a los proveedores de imágenes a subir imágenes digitales de diferentes estilos de pintura, como el cubismo, el impresionismo, el realismo, diferentes paisajes (campo, naturaleza muerta, retratos, etc.) o fotografías personales. Una vez almacenadas en el sistema, se seleccionaron algunas de ellas con el objetivo para representar una muestra de los diversos estilos de pintura que los usuarios han decidido utilizar para nuestro estudio. La Figura 7.7 (a) es una copia digital de “La Persistencia de la Memoria” de Dalí y representa el arte surrealista. La Figura 7.7 (b) es una fotografía con una vista del patio del edificio histórico de la Universidad de Salamanca. La Figura 7.7 (c) es una fotografía de una de las mascotas de los usuarios que han participado en el sistema. La Figura 7.7 (d) y (e) son fotografías de paisajes naturales. La Figura 7.7 (f) se titula “La Creación” de Miguel Ángel. La Figura 7.7 (g) “El Grito” de Edvard Munch, mientras que la Figura 7.7 (h) es una copia digital de la “Noche Estrellada” de Van Gogh.

<sup>3</sup><https://form.jotformeu.com/70101890150342>

TABLA 7.2: Las estadísticas resultantes de la organización de las calificaciones subjetivas en función de las medidas de distancia de la representación TIS.

$\chi^2$	$M_e$	$M_o$
$6.8381e - 04$	Bueno	Bueno

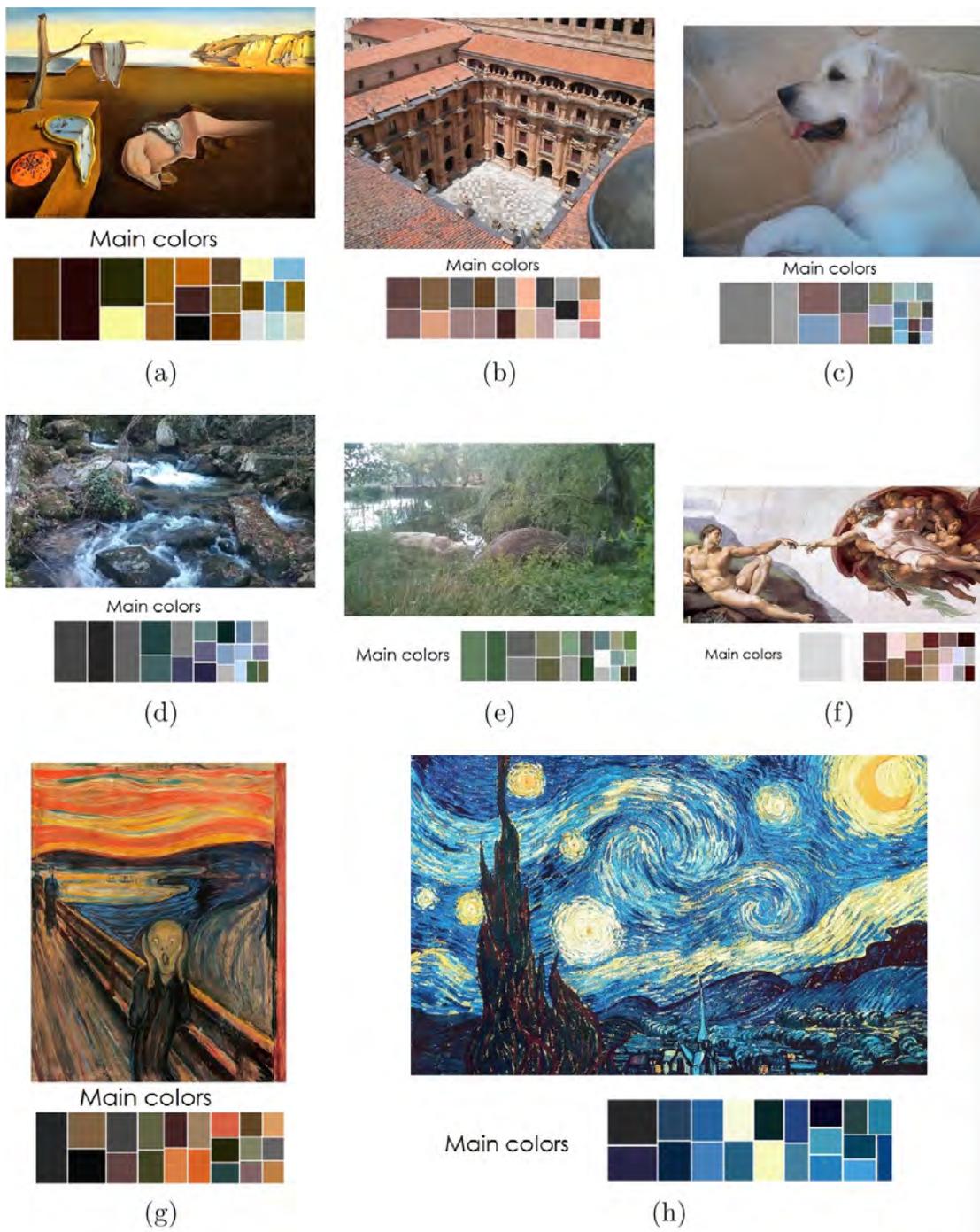


FIGURA 7.7: Colección de imágenes aplicadas al sistema para extraer música de nube de sonido.

En este caso, se consideraron 42 expertos musicales que calificaron la percepción de similitud entre música y sonido. La Figura 7.8 muestra los resultados finales obtenidos.

La gráfica muestra el nivel de similitud entre la melodía y el color considerado por los oyentes, así como la desviación estándar obtenida para cada fragmento musical. Los valores también se normalizaron con fines de clarificación. Todas las tasas medias

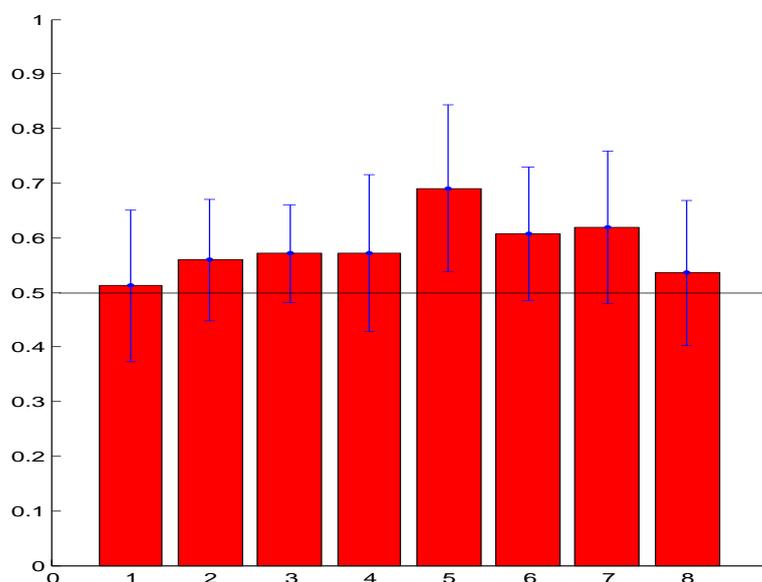


FIGURA 7.8: Muestra las tasas de similitud entre los colores y el sonido según los oyentes.

están por encima de 0.5 (“algo similar”) según las tasas subjetivas. Sin embargo, la desviación estándar calculada muestra que algunas opiniones están por debajo de este umbral, considerando la melodía como no suficientemente similar. Cabe señalar que la percepción de la similitud entre los colores y el sonido es muy personal y depende de factores como la cultura o el estado de ánimo. Por lo tanto, podríamos haber obtenido resultados muy diferentes para el mismo fragmento musical dependiendo del contexto del oyente. Además, decir a los expertos que hay una conexión entre la imagen y la música puede sesgar el análisis. Proponemos como trabajo futuro desarrollar un nuevo análisis para comparar las puntuaciones recibidas por melodías asociadas con imágenes y melodías que no lo son.

En general, la melodía apunta a seguir los principales estándares tonales pero siempre está inspirada en una pintura. Por lo tanto, la melodía puede no ser una pieza tonal completa, no se designa para ser una melodía clásica. Aunque los estándares tonales occidentales son bastante convencionales cuando se les pide evaluar una melodía, la música es siempre un arte subjetivo que depende de las preferencias personales y la cultura.

Finalmente, la similitud entre el color y los sonidos fue muy difícil de medir, ya que las percepciones personales desempeñan un papel crucial en los resultados finales. Este tema necesita un análisis más profundo en el que se deben considerar las cuestiones psicológicas, sociales y culturales.

## 7.5 Resultados del Componente de Aprendizaje

En este nuevo módulo se integra el ciclo CBR que se definió en el Capítulo 6. En este caso el objetivo es conocer si esta nueva incorporación supone verdaderamente una mejora en el framework total. Debido a la naturaleza de los resultados finales, hemos realizado un segundo experimento con expertos que han validado este nuevo módulo.

En la Figura 7.9 se presentan seis imágenes diferentes, que fueron seleccionados para preservar la variedad y demostrar que el algoritmo trabaja con imágenes extraídas de diferentes contextos y estilos artísticos.

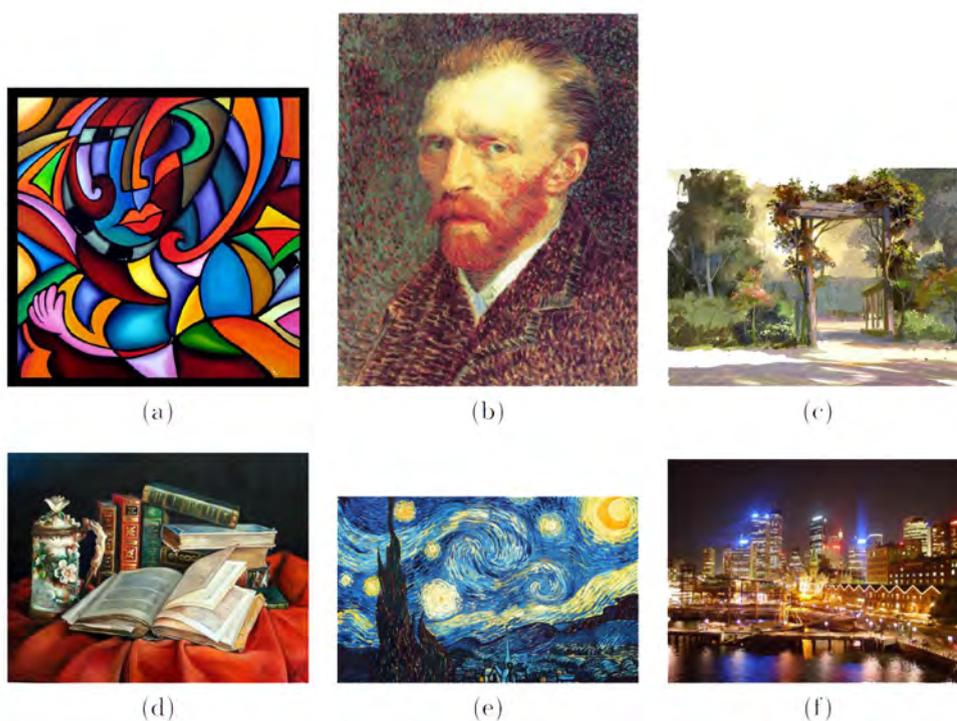


FIGURA 7.9: Colección de imágenes utilizadas en el experimento.

Se seleccionaron 57 expertos en música para probar dos estudios diferentes. El primer análisis consistió en calificar el fragmento musical obtenido de las imágenes. Así, se presentó a los expertos el fragmento musical y el cuadro digital en la misma prueba. Al igual que antes, tuvieron que evaluar de muy bueno a muy malo, no sólo la calidad de la composición musical, sino también si la composición musical podía describir los colores de la imagen correctamente de acuerdo a sus percepciones. Los resultados finales se representan en la Figura 7.10.

El eje horizontal del gráfico representa el número de la imagen y el eje vertical la puntuación global que las imágenes han obtenido. Como podemos ver, la imagen correspondiente a la opción c fue muy bien calificada. Todas las muestras tienen además

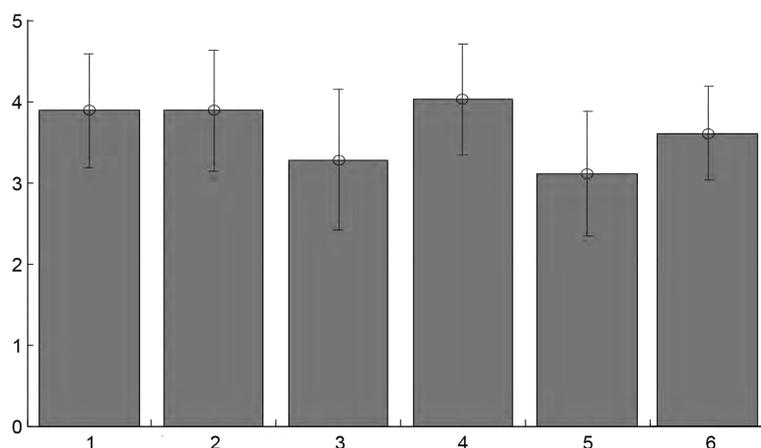


FIGURA 7.10: Resultados de la evaluación para cada fragmento extraído de una imagen. Cada número corresponde a una imagen de la Figura 7.9 en el mismo orden.

un valor medio por encima de “Regular”. Sin embargo, como son puntuaciones muy subjetivas, los resultados pueden diferir fuertemente de algunos usuarios a otros. Por ello, también se representa con las tasas de error visualizadas como barras para cada caso.

Nuestro objetivo es mejorar los resultados obtenidos en la sección anterior. Ambos contienen un componente social capaz de generar productos creativos y que permite la interacción entre los usuarios. Sin embargo, el sistema anterior no proporciona ninguna regeneración que permita considerar la opinión social para mejorar la composición final de la música. Por lo tanto, hemos añadido una arquitectura CBR integrada en el rol Planificador para aprender de las experiencias anteriores teniendo en cuenta la valoración de los expertos y así proporcionar melodías mejoradas. Esta mejora con el componente de aprendizaje buscaba fomentar el propósito de la computación social, como la interacción social y la generación de resultados creativos, utilizando el contexto social como componente de retroalimentación.

Finalmente se realizó un estudio comparativo entre los resultados obtenidos del sistema con y sin la arquitectura CBR, que representa el segundo caso de estudio. En esta ocasión, se mostró a los expertos la misma imagen y dos fragmentos musicales. La pregunta a responder era qué fragmento pensaban que sonaba mejor de acuerdo con los términos de calidad y la adaptación a la imagen. Se presentaron las mismas seis imágenes con dos fragmentos, el fragmento A que representa el fragmento obtenido con la arquitectura CBR y el fragmento B, que representa el resultado obtenido por el sistema sin el componente de aprendizaje. Tuvieron que seleccionar el mejor fragmento según criterios de calidad y adaptación a la imagen según su percepción. Podrían seleccionar

Fragmento A (barras azules) como el mejor, Fragmento B (barras verdes) o Indeciso (barras rojas) si pensaban que ambos fragmentos eran de calidad similar. La Figura 7.11 muestra los resultados finales con esta segunda forma.

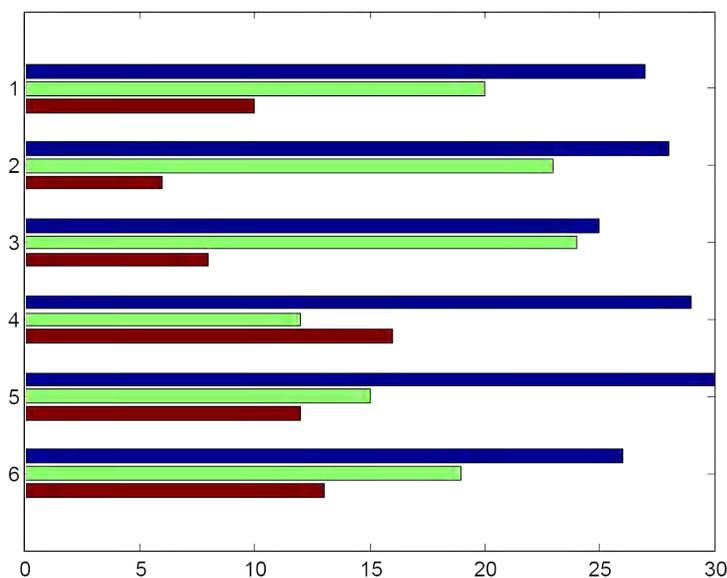


FIGURA 7.11: Comparación entre las validaciones del sistema sin el componente de aprendizaje y el actual. La votación para el componente CBR como el mejor resultado se destaca en azul, la votación para el trabajo sin el aprendizaje se resalta en verde, y finalmente el voto indeciso se representa con una barra roja.

En el eje horizontal hay seis grupos correspondientes a las seis imágenes analizadas con los fragmentos. Las barras azules representan el número de veces que los fragmentos A con la arquitectura CBR implicada fueron seleccionados como los mejores. Las barras rojas indican el número de veces que el fragmento B se selecciona como el mejor. Las barras verdes recogen el número de veces que se ha seleccionado la opción “Indefinido”. Como podemos ver, el fragmento A se selecciona principalmente como la mejor opción en cada caso. Así, podemos concluir que la arquitectura CBR que incorpora la opinión social para aprender de experiencias anteriores, permite mejorar los resultados del sistema social. El algoritmo ACO permite navegar por el espacio de búsqueda y seleccionar el mejor camino a través de las notas musicales para extraer la mejor melodía de los casos anteriores.

Cabe destacar que, al igual que pasaba en el caso anterior, la composición del sonido es una melodía que sigue algunos estándares de música tonal, aunque no puede considerarse una composición tonal completa. Nuestro objetivo es expresar una obra de arte pictórica a través de otra ciencia artística como la música. Además, como los experimentos sólo son llevados a cabo por expertos con formación musical específica, los resultados pueden ser sesgados debido a esta educación musical. Proponemos ampliar este trabajo

a personas con diferentes formaciones musicales con el fin de proporcionar un análisis más profundo de la calidad musical y la aceptación social.

## 7.6 Evaluando el Framework

El framework descrito en este documento se implementó como una red social y se probó durante varias semanas, donde el componente humano incluye a los miembros de la red social, y el componente máquina está compuesto por la organización virtual. Este prototipo de organización virtual ha sido ajustado y actualizado durante este período y los resultados iniciales han sido positivos desde un punto de vista técnico y científico. La construcción del prototipo fue relativamente fácil de llevar a cabo utilizando el análisis y diseño obtenidos aplicando la metodología GORMAS, lo que ha ayudado positivamente a obtener resultados técnicos en función de los problemas de eficiencia y flexibilidad. En particular, el rol Planificador fue diseñado para implementar una arquitectura CBR, que es capaz de aprender de experiencias pasadas y es esencial para el proceso de toma de decisiones. Por tanto, el agente Planificador representa el núcleo de aprendizaje del framework. Otro punto esencial es el Agente Interfaz, que funciona como nexo de unión entre los diferentes módulos para formar un todo, que es nuestro framework.

Los datos derivados de las interacciones entre humano y máquina son manipulados automáticamente por la organización durante todo el proceso, extrayendo la información necesaria para generar un nuevo caso a partir de los colores de entrada, o para ir recibiendo los datos del usuario en el caso de la composición de acordes. Cada evaluación del usuario obtenida a través del framework es analizada y almacenada en la base de datos.

El framework se desplegó finalmente como una red con más de 80 miembros voluntarios. Los usuarios podían ser cualquier tipo de individuos, pero los expertos para dar su opinión fueron seleccionados de acuerdo a su educación musical (por lo menos 6 años de experiencia musical o de formación), para obtener resultados más útiles. En particular, este framework fue probado durante 25 días. 83 usuarios fueron inscritos como miembros del sistema, algunos de ellos con formación profesional en conocimiento musical.

También presentamos una comparación entre tres sistemas para resaltar las ventajas de nuestro trabajo. En particular, se estudia un trabajo creativo basado en MAS [Kirke and Miranda, 2015] junto con el sistema JANUS [Gaud et al., 2008], un marco que trabaja con organizaciones virtuales y MAS para propósitos generales. La comparación cualitativa se muestra en la Tabla 7.3.

	Nuestro Framework	JANUS	Propuesta de Kirke y Miranda
Productos creativos y X	-	X	
Interacción	X	X	X
Sistema Abierto	X	X	-
Capacidad de Crecimiento	X	X	-
Carácter Social	X	X	-
Participación Pública	X	X	-
Utiliza un MAS	X	X	X
Arquitectura BDI	X	X	X
Computación en Enjambre	X	-	-
Combina el razonamiento con otros algoritmos	X	-	-
Soporte para VOs	X	X	-
Compatible con servicios web	X	X	-
Ejecutable en diferentes SO	X	X	-
Apoyo a expertos para interactuar con el sistema	X	X	X
Balance de carga	X	X	-
Proporciona una interfaz de usuario	X	-	X
Proporciona una herramienta de registro	-	X	-
Implementa el concepto de holón	-	X	-

TABLA 7.3: Comparación de nuestro framework, el sistema JANUS y el trabajo propuesto por Miranda & Kirke'

Vale la pena señalar que el estudio desarrollado por [Kirke and Miranda, 2015] utiliza emociones y MAS para generar una melodía musical. El uso de emociones para crear nueva música realza la originalidad del sistema. Sin embargo, aunque interactúa con los usuarios, no es una aplicación adecuada para las comunidades sociales. Además, la escalabilidad y la flexibilidad no están incluidas en su trabajo, ya que no diseñaron el sistema siguiendo una metodología basada en Organizaciones Virtuales.

JANUS [Gaud et al., 2008] es una plataforma multiagente diseñada específicamente para manejar la implementación y despliegue de sistemas multiagente. Se basa en un enfoque organizativo y se centra en apoyar la implementación de los conceptos de rol y organizaciones como entidades de primera clase. Esta consideración tiene un impacto significativo en la implementación del agente y permite a un agente cambiar de forma fácil y dinámica su comportamiento. Esta característica también se comparte con

nuestro sistema, apoyando el diseño de las VOs junto con la composición musical. La plataforma JANUS también contribuye a llenar la brecha entre las fases de concepción e implementación en este ámbito. El sistema también implementa el concepto del agente holón. Un holón se define como una entidad autónoma que puede ser modelado recursivamente con sub-componentes de un sistema más grande [Gaud et al., 2008]. Aunque este concepto facilita la gestión de la jerarquía de agentes y el desarrollo de MAS, la implementación de agentes holónicos en nuestro sistema queda fuera del alcance del trabajo presentado aquí. Además, nuestro sistema comparte la mayoría de las características correspondientes a un marco de organización virtual, agregando un componente social esencial para el éxito de un sistema de composición musical.

# Capítulo 8

---

## Conclusiones

---



**VNiVERSIDAD  
D SALAMANCA**



# Conclusiones

---

En este trabajo se ha propuesto un framework que ha integrado una organización que permite la interacción entre agentes virtuales y agentes humanos para generar música. El resultado musical ha seguido en la medida de lo posible las preferencias de los distintos usuarios que han participado de nuestro framework. Recordemos que la generación de un resultado musical supone un desafío debido a la naturaleza de la música, un producto creativo que depende de factores subjetivos como son la cultura o el estado de ánimo de los usuarios, entre otros. El primer prototipo se despliega como una red social con casi 100 miembros que deben proporcionar las entradas digitales y evaluar los resultados musicales. Teniendo en cuenta los resultados obtenidos, podemos decir que el framework ha sido capaz de generar música generalmente aceptable para una muestra de la sociedad que hemos seleccionado para nuestro estudio.

El framework se ha dividido en dos partes principales. La primera de ellas contiene un agente capaz de generar una progresión de acordes siguiendo normas relativas a la tonalidad. La segunda contiene una organización de agentes que genera melodías a partir de una imagen digital proporcionada por los usuarios. Ambas partes necesitan del usuario para poder crear bien progresiones de acordes, bien melodías. En el primer caso, el sistema genera varias opciones para cada acorde que se quiere añadir a la progresión. El usuario tiene siempre la última palabra, puesto que debe seleccionar qué acorde debe incorporarse al resto de la progresión de entre todas las opciones presentadas. En el segundo caso, el sistema recoge las valoraciones de los usuarios para mejorar los resultados musicales finales. Para ello, se desarrolla e integra en nuestra sociedad una arquitectura de Razonamiento Basado en Casos (CBR). Así, tenemos dos roles humanos: los proveedores sociales que seleccionan los acordes para incorporarlos a la progresión, o bien dan las imágenes al sistema para extraer el color, y los expertos sociales, que evalúan la calidad de la música generada.

Por tanto, hemos construido una sociedad humano-agente, ya que el ser humano no solo proporciona al sistema la entrada de datos necesaria para la generación musical, sino que es un colaborador activo capaz de influir en el resultado final mediante la evaluación

y/o la corrección de los mismos. Los HAS no se han aplicado aún a un contexto creativo como la composición musical. Esto fue parte del detonante para añadir un componente humano y a experimentar con una nueva forma de generar música basada en contextos sociales. Por tanto, esta sociedad humano-agente capaz de generar contenido creativo de forma dinámica puede considerarse como la contribución principal de esta tesis al estado de la técnica.

La computación social ha sido aplicada con éxito para crear interacción entre expertos y nuevos productos basados en opiniones sociales, considerando este punto como una nueva contribución al campo de la aplicación de computación social a contextos creativos. En nuestra propuesta actual, el componente social es esencial en el proceso de progresión de acordes, para elegir cuáles de ellos se incorporarn a la progresión, y en el proceso de producción de melodías para transformar la imagen en música, ya que motiva el inicio del proceso de cargar una obra de arte pictórica y porque la opinión social se tiene en cuenta para mejorar los resultados futuros.

Para codificar y medir la calidad de las distintas configuraciones tonales, se ha utilizado un espacio geométrico, denominado “Tonal Interval Space” (TIS). El espacio TIS es capaz de representar diferentes configuraciones tonales tales como notas, intervalos, acordes y escalas de una forma similar. El TIS demostró ser una representación adecuada por múltiples razones. En primer lugar, es capaz de representar configuraciones tonales multinivel. Las propiedades topológicas del espacio reflejan las propiedades musicales del sistema tonal. En consecuencia, podemos proponer una función objetiva que codifica reglas sobre progresiones de acordes o sobre melodías como medidas geométricas entre diferentes niveles de configuración tonal, tales como acordes y escalas. Además, el TIS es capaz de manejar diferentes escalas en diferentes modos, permitiendo la generación de progresiones de acordes o melodías en cualquier tonalidad con el mismo procedimiento. Aunque este espacio ya había sido desarrollado por Bernardes et al. [2016], no se había utilizado hasta ahora para codificar medidas de calidad de configuraciones tonales, ni para generar un sistema capaz de componer melodías y progresiones de acordes.

En el caso de la generación de progresiones, a partir de una secuencia de acordes de entrada, el agente puede buscar el acorde siguiente siguiendo restricciones musicales tales como consonancia, tonalidad y función armónica codificadas como medidas geométricas en una nueva función objetivo que está siendo optimizada. En esta parte, la contribución principal consistió en adaptar el algoritmo de Sistema Inmune Artificial llamado opt-aiNet e integrarlo en un framework para encontrar múltiples soluciones en paralelo, dando como resultado múltiples acordes candidatos para seguir la secuencia de entrada. A partir de ahí, el usuario podía seleccionar qué acorde debía incorporarse a la progresión de entre estas opciones presentadas. Se realizó una prueba de audición para evaluar la

calidad subjetiva de las soluciones candidatas propuestas por opt-aiNet. La mayoría de los oyentes calificaron los acordes propuestos por el AIS como mejores candidatos para seguir las secuencias de entrada que los acordes descartados por el AIS. El análisis estadístico mostró que las valoraciones subjetivas se correlacionan fuertemente con los valores objetivos, validando la función objetivo como un proxy para la medida subjetiva de la calidad. También hemos determinado que la distancia euclídea al acorde anterior utilizando múltiples representaciones diferentes no es un buen indicador de la calidad subjetiva. Estos resultados demuestran que la contribución, tanto en el campo de la evaluación de la calidad de los acordes a partir de medidas matemáticas, así como en la adaptación de un algoritmo AIS para generar configuraciones tonales (acordes), ha sido positiva de acuerdo con los usuarios del sistema.

En la prueba de audición, la evaluación de un acorde en una progresión depende del contexto musical. En otras palabras, la evaluación de las progresiones de los acordes depende no sólo del acorde anterior, sino también del siguiente en la secuencia. El mismo candidato puede considerarse una buena elección o una mala dependiendo del siguiente acorde en la secuencia. Por ejemplo, uno disonante puede ser considerado mejor cuando es seguido por la tónica, porque la tónica proporcionaría una resolución a la tensión provocada por el acorde disonante. Musicalmente, la evaluación subjetiva de las progresiones depende también de factores como la cultura, el estado de ánimo, la educación musical o la preferencia personal que no están codificados en la función objetivo. En consecuencia, los resultados de la prueba de audición varían considerablemente entre los individuos, reflejando la naturaleza subjetiva de la tarea. Por ejemplo, las personas con formación musical pueden ser más tolerantes a la disonancia debido a la exposición a la música artística o a la música de otras tradiciones que no siguen el sistema tonal occidental. En este caso, un acorde disonante podría ser interpretado como un cambio estético en lugar de una mala elección. Pese a ello, se obtuvieron unos resultados muy positivos que nos animan a seguir investigando en esta nueva forma de generar progresiones de acordes.

En el caso de la generación de melodías, la organización virtual contiene una serie de agentes dotados de un mecanismo de evaluación de calidad utilizando el espacio TIS. Estos agentes, llamados Agentes Sinestésicos, miden la calidad de una nota basándose en notas anteriores y la tonalidad. Este mecanismo es esencial para generar una melodía que siga los cánones tonales. Este sistema de generación melódica se diseñó con una organización virtual e implementó con un sistema multiagente para crear un framework flexible que pueda adaptarse a nuevos contextos dados por diferentes situaciones sociales.

La organización sigue un flujo de trabajo de cuatro etapas. El primer paso consiste en la extracción de color de la imagen proporcionada por la comunidad social. Las

propiedades del color se usan para dar los parámetros iniciales a los Agentes Sinestésicos, quienes a su vez valoran la calidad del sonido generado para moverse por el espacio. El movimiento originado en el espacio sigue diversas reglas de acuerdo con un algoritmo de enjambre modificado y diseñado para este trabajo en particular. Con el fin de generar una composición musical consistente, se desarrolla un agente para evaluar la probabilidad de cada sonido basado en los sonidos previos existentes en la composición. Este resultado final se sintetiza y se reproduce para una comunidad de expertos.

Para evaluar la aceptación de la música creada, se recurrió a las opiniones de expertos musicales. En concreto, se diseñaron dos pruebas de audición con los mismos fragmentos musicales, uno para medir la calidad musical, y el otro para recoger la percepción general de los expertos sobre la similitud de los colores y el sonido. Ambas pruebas dieron resultados alentadores: la calidad de la música era generalmente bien clasificada y la asociación entre música y sonido era bastante bien percibida entre nuestros usuarios. La contribución principal en la generación de melodías es aplicar una organización virtual para obtener un producto creativo que depende el contexto, y en la relación establecida entre dos artes perceptualmente diferentes, como son la imagen y el sonido.

Para desarrollar el sistema de aprendizaje, se modificó el módulo de extracción de color, y se aplicó un método de extracción y comparación de histogramas de color en lugar de guardar la codificación de nuestros colores en bruto. Estos histogramas almacenados son utilizados para seleccionar experiencias anteriores con características similares. Estas experiencias se llaman casos, y constituyen la memoria del sistema. Los casos recuperados se adaptan utilizando el algoritmo de hormigas ACO para seleccionar aquellos caminos que unen las mejores notas musicales y crear un nuevo caso basado en los anteriores. Este proceso de selección considera la opinión social de los expertos musicales, así como la calidad de la música asociada a cada nota individual basada también en el espacio TIS, que mide propiedades como la consonancia, la distancia entre notas o la tonalidad. El resultado final es sintetizado para la comunidad social, que se encarga de evaluar los resultados según criterios musicales y percepción de similitud entre música y colores. Estas valoraciones se utilizan para decidir si el resultado musical se puede utilizar en el futuro. El proceso principal para recuperar casos similares, adaptarlos y decidir si se pueden almacenar en la memoria es llevado a cabo por una entidad específica, el Planificador, que se construye siguiendo una arquitectura CBR. La contribución principal en este caso ha sido el poder integrar un aprendizaje en la computación social, basándonos en las opiniones de los usuarios para generar un contenido nuevo tan dinámico y dependiente de la subjetividad de los usuarios.

Los resultados finales fueron filtrados y se seleccionaron varias imágenes con su correspondiente composición musical, que fueron evaluados por expertos y comparados

con los resultados obtenidos en el trabajo anterior en términos de calidad musical y adaptación a los colores de la imagen. A pesar de la posible subjetividad de las valoraciones, los resultados globales indican que la calidad puede ser aceptable para generar música a partir de colores. Las puntuaciones proporcionadas por los expertos nos permiten concluir que nuestro enfoque actual mejora sus resultados aprendiendo del contexto social, principal contribución de este sistema.

A continuación, se describirán en mayor detalle las contribuciones principales al campo de la investigación que nos ocupa.

## 8.1 Contribución al Estado de la Técnica

En este trabajo se ha propuesto un framework que integra organizaciones virtuales para generar resultados musicales, utilizando la interacción humano-agente. Por tanto, esta tesis realiza algunas contribuciones en el ámbito de la computación social y en la generación automática de música. Las aportaciones principales se resumen como sigue:

- **Generación de música a través de mecanismos evolutivos.** El comportamiento de los agentes se ha adaptado a las características de la organización para componer música. En particular, los agentes encargados de la composición musical han requerido la especificación de idiomas para definir reglas y motores capaces de interpretarlos en tiempo de ejecución. En este caso, la representación utilizada para evaluar las características de una composición musical, ha sido el espacio TIS. Este espacio se ha adaptado fácilmente a los algoritmos evolutivos, lo cual ha permitido la elaboración de reglas de manera colaborativa entre los agentes de una organización, facilitando un comportamiento distribuido y adaptativo a las circunstancias de las organizaciones, tales como la transformación de color en sonido, o la extracción de varios acordes candidatos de forma paralela. En concreto, las contribuciones han sido:
  - Revisión y estudio de diferentes modelos y sistemas de composición utilizando sistemas evolutivos.
  - Definición de nuevas medidas de evaluación musical basadas en propiedades geométricas de un espacio, tales como la distancia euclídea o el ángulo.
  - Adaptación de un sistema inmunológico (opt-Ainet) a la composición musical de acordes.
  - Adaptación de un algoritmo en enjambre (PSO) a la generación de melodías utilizando colores como fuente de inspiración.

- Adaptación de un algoritmo ACO para seleccionar una nueva melodía entre las ya conocidas, consideradas como “experiencias previas”.
- **Técnicas para el aprendizaje social.** La definición de una sociedad humano-agente requiere mecanismos que faciliten el aprendizaje social de los comportamientos humanos para reproducirlos en las sociedades de agentes. Las contribuciones en este campo han sido:
  - Revisión del estado del arte de las técnicas que permiten un enlace entre los estímulos y las respuestas hechas por los seres humanos, así como el uso de sociedades dentro de la composición musical.
  - Aplicación de una arquitectura CBR para almacenar y reutilizar composiciones musicales con el fin de mejorar los resultados de nuestra propuesta.
  - Incorporación del comportamiento humano en nuestro sistema mediante la valoración y/o corrección de los resultados para mejorar posteriores composiciones musicales. En nuestra propuesta, el usuario es un colaborador activo, sin él es imposible producir un resultado musical.
- **Descripción e implementación de la simulación.** Se ha generado una simulación donde los usuarios (el componente humano) han interactuado y evaluado el conjunto de nuestro framework. Este primer prototipo era una red en la que los usuarios colaboraban con la máquina para extraer y evaluar una composición musical. Se ha utilizado un sistema CBR para extraer información y patrones de comportamiento. Este conocimiento puede ser útil para futuras decisiones de generación musical. En concreto, se ha contribuido de la siguiente manera:
  - Se ha dotado a los agentes compositores de cierto comportamiento humano creativo, esencial para la generación de música.
  - Se ha realizado una evaluación exhaustiva del framework basándonos en la experiencia humana y en la evaluación de los resultados por parte de los expertos.
  - Se ha modelado de un sistema automático de extracción de información y mecanismos de análisis para mejorar las generaciones musicales.

Finalmente, cabe destacar que se ha realizado un estudio del estado del arte en sistemas creativos capaces de generar música siguiendo diversas técnicas, pero profundizando en aquellos que empleaban sistemas bio-inspirados y/o sistemas multiagente. Asimismo, se han estudiado los sistemas y organizaciones de agentes. También se han analizado las

diferentes técnicas evolutivas que han sido empleadas en nuestro trabajo. Se ha hecho especial hincapié en el nuevo concepto de computación social, esencial para la definición y diseño de una sociedad humano-agente como la desarrollada en este trabajo. A partir de estos estudios, se ha procedido a construir un framework basado en organización de agentes que incorpora rasgos sociales y modelos de aprendizaje para generar música colaborando con los humanos, que son los usuarios.

## 8.2 Trabajo Futuro

Debido a la multidisciplinariedad de esta tesis, existen una gran cantidad de puertas que dan pie a un análisis y trabajo futuros.

En relación a nuestro sistema de representación, una limitación del espacio TIS surge del vector croma como el núcleo de la representación. El vector croma colapsa todas las octavas en una, eliminando la posibilidad de representar clases tonales en diferentes octavas, y consecuentemente acordes con inversiones. Representar acordes con inversiones en el TIS sigue siendo un desafío, que puede proponerse como trabajo futuro.

En este trabajo nos centramos en el aspecto a corto plazo de la generación automática de progresiones de acordes. Creamos un sistema capaz de realizar una búsqueda del siguiente acorde en una secuencia, dados dos acordes anteriores como entrada. En particular, nos centramos en la capacidad de proporcionar múltiples opciones que se consideran buenas soluciones candidatas que son diferentes entre sí. No obstante, las progresiones más largas requieren una medida que recoja más de dos acordes en una secuencia, y estudie la calidad del contexto musical. El trabajo futuro debería investigar cómo generar automáticamente progresiones más largas, mediante el diseño de alguna medida que capture la calidad de un acorde de acuerdo al resto de la progresión. Adicionalmente, se podría utilizar la valoración de los expertos para añadir un mecanismo de aprendizaje basado en experiencias pasadas.

Nuestro sistema generador de melodías es capaz de crear música aceptable para los usuarios. Sin embargo, el número de imágenes y melodías utilizados en nuestro análisis podría incrementarse para extraer conclusiones aún más sólidas. Asimismo, las percepciones de la correlación entre colores y sonidos deben investigarse más a fondo con factores psicológicos para demostrar que la transformación color-sonido es efectiva desde un punto de vista social. Por tanto, se pretende ampliar nuestro análisis, no solo a expertos en música, sino también a un amplio contexto social con personas con diferente formación musical, para analizar la aceptación social de nuestros resultados musicales.

La transformación de imágenes en notas es algo polémico y subjetivo, que depende no sólo del individuo, sino también del estado de ánimo del individuo y de otros factores. En un estudio futuro, desarrollaremos una nueva prueba para mejorar cómo se recoge la percepción de la correlación entre la música y el color. Además, se propone una mejora del análisis de imagen para recoger no solo colores, sino también formas y disposición de figuras para adaptarnos mejor al contexto pictórico. Otro problema detectado es que el número de colores determina las notas que se utilizarán en la composición. Cuando una imagen tiene colores muy uniformes, la melodía contiene una baja diversidad de notas. Para trabajos futuros, proponemos extraer motivos musicales utilizando las notas dadas en el sistema, y utilizarlos como base para una composición musical siguiendo otros parámetros que también podemos extraer de la imagen, como formas o distribución de elementos.

Sería también interesante probar otro tipo de algoritmos bioinspirados, como los Sistemas Inmune o Algoritmos Genéticos, para hacer nuevos casos de los anteriores usando operadores de cruce y mutación.

Hasta ahora, todos los usuarios son considerados como iguales dentro de la sociedad, y su opinión es válida por igual, independientemente del rol que desempeñe en nuestra sociedad. Podría darse el caso que un usuario fuera especialmente exigente o bien especialmente laxo. En ambos casos, sus valoraciones deben estudiarse de forma diferente a la de un usuario medio. Un futuro análisis pretende incorporar mecanismos de reputación, muy utilizados en computación social para medir el alcance que tiene un usuario en la comunidad y su fiabilidad en la evaluación.

# Apéndice A

---

## Una Breve Introducción a la Teoría Musical

---



VNiVERSIDAD  
DSALAMANCA



# Una Breve Introducción a la Teoría Musical

---

## A.1 Introducción

A la hora de leer esta tesis, es importante tener en cuenta algunos conceptos de la teoría musical. Esto nos permitirá el entendimiento de algunos parámetros musicales que aquí se aplican. Además, tendremos a nuestro alcance el poder analizar la generación de melodías y armonías desde un punto de vista musical, no solo computacional.

La música está basada eminentemente en sonidos, compuestos a su vez por las ondas sonoras que atraviesan el medio y llegan al oído. De acuerdo con la teoría musical, todo sonido tiene cuatro características: altura, intensidad, duración y timbre.

- **Altura o Tono:** La altura o tono está relacionada con la magnitud física de la frecuencia de onda, es decir, el número de vibraciones que una onda realiza por segundo. En función de esta característica, los sonidos pueden dividirse en graves y agudos. Cuanto mayor sea la frecuencia, más agudo será un sonido; cuanto menor frecuencia, más grave.
- **Intensidad:** La intensidad está relacionada con la magnitud de amplitud de onda. De acuerdo con esta magnitud, los sonidos pueden dividirse en fuertes, si la amplitud de onda es muy grande, o en suaves o piano, cuando la amplitud de onda es pequeña.
- **Duración:** La duración se refiere al tiempo que dura la ejecución de un sonido, pudiendo por tanto encontrar sonidos largos o cortos.
- **Timbre:** El timbre tiene que ver con la forma de la onda, lo que nos permite distinguir o identificar sonidos provenientes de diferentes fuentes. De este modo, podemos diferenciar un sonido producido por un piano, un violín o una voz humana.

Cuando los sonidos se ordenan siguiendo una serie de patrones que pueden estar predefinidos o no, estamos ante una composición musical. Toda composición contiene varios elementos:

- **Melodía:** Una melodía es una organización de alturas de sonidos de forma horizontal. De este modo, en una melodía nunca se superponen más de un sonido al mismo tiempo.
- **Armonía:** La armonía es una organización de alturas de los sonidos, pero de forma vertical. En este caso, varios sonidos pueden superponerse de forma simultánea, por ejemplo cuando varias personas cantan al mismo tiempo diferentes notas, o tocan diversas notas en sus instrumentos, o cuando se tocan varias notas de un instrumento polifónico como el piano. La armonía se relaciona con la organización de las alturas de los sonidos en un sentido vertical.
- **Ritmo:** Se refiere a la distribución de la duración de los sonidos en el tiempo. Cuando escuchamos una pieza musical percibimos una sensación de movimiento. Esa sensación es el resultado de la forma en que las distintas notas se han distribuido en la pieza.
- **Métrica:** La métrica se refiere a patrones regulares en el ritmo que se hacen evidentes a través de los acentos naturales.
- **Dinámica:** Hace referencia a la combinación de intensidades de los sonidos.

Los diferentes sonidos que se utilizan en música reciben diferentes nombres o notas. Estas notas tienen la particularidad de estar asociadas a frecuencias de onda fijas, y cada frecuencia recibe un nombre y una grafía, que depende de la clave en la que estemos. Una clave no es más que un punto de referencia para poder nombrar las diferentes notas en una partitura. En nuestro caso, y para mayor brevedad, solo veremos la clave de Sol como punto de referencia para nombrar las notas.

Para expresar gráficamente una composición musical, ya sea melodía, armonía o ritmo, se utilizan los pentagramas. Un pentagrama es un conjunto de cinco líneas y cuatro espacios, paralelas y equidistantes, tal como se ilustran en la Figura A.1.

Los sonidos más graves se escriben en la parte más baja del pentagrama. Conforme un sonido musical se va haciendo más agudo, su nota va ascendiendo. Puede ocurrir que las notas sean demasiado agudas o demasiado graves para poder representarse solo con las 5 líneas del pentagrama. En este caso, podemos añadir más líneas a través de lo que se conoce como líneas adicionales (Figura A.1).

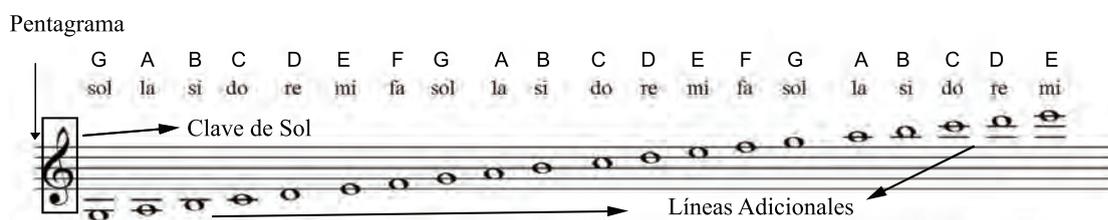


FIGURA A.1: Representación de una serie de notas colocadas de forma ascendente. Las notas que se encuentran en la imagen llevan el nombre justo encima siguiendo las sílabas musicales y el alfabeto musical.

A lo largo de la historia se han desarrollado diferentes formas para nombrar las notas musicales. En este trabajo y actualmente, se reconocen dos formas de denominación:

- El Alfabeto Musical, basado en las primeras siete letras del alfabeto, que deriva de la forma en la que llamaban a las notas en la antigua Grecia. El uso de este sistema está más extendido en Inglaterra y Estados Unidos, por eso, algunos lo llaman Notación Anglosajona.
- Las sílabas musicales, utilizadas en el solfeo. Guido de Arezzo creó otro sistema en el que las notas fuesen más fáciles de cantar y entonar. Los nombres de las notas provienen de la primera sílaba de cada verso en la letra de una canción. Además, cada sílaba en la melodía está en la nota que corresponde con el nombre designado. Con este sistema es fácil recordar los nombres de las notas y su entonación correspondiente, por eso se denomina sistema mnemotécnico.

La equivalencia entre los dos sistemas de denominación se representan en la Figura A.1.

Como vemos, tradicionalmente se generó el nombre para 7 notas musicales. Sin embargo, cuando escuchamos una composición musical, suele haber más de siete sonidos. Existen dos formas principales para representar el resto de sonidos, basándonos en la nomenclatura de estas 7 notas musicales: las alteraciones, que explicaremos un poco más adelante, y el concepto de octava.

El concepto de octava es de gran importancia en el desarrollo de nuestro sistema musical. Consiste en reevaluar las notas con respecto a la altura musical del sonido. En este sentido, si seguimos en orden las notas, llegaremos a un punto en que repetiremos la misma nota pero a una altura superior. En este caso decimos que la nota está una octava más alta.

En la música occidental tradicional, la octava se divide en 12 sonidos que están equidistantes entre sí. Estos doce sonidos se obtienen mediante las siete notas musicales y estas utilizando las alteraciones que veremos un poco más adelante.

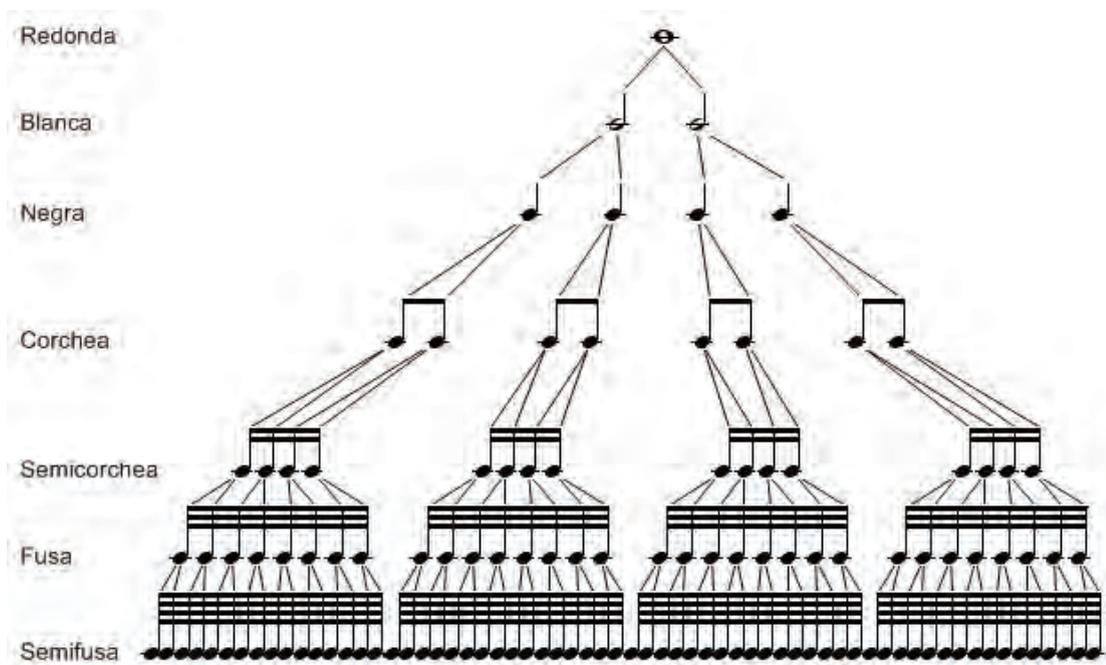


FIGURA A.2: Representación la relación de las diferentes figuras rítmicas que tienen entre sí.

Dentro de la teoría musical, existen varios elementos que se pueden emplear para crear una composición. Estos elementos se pueden denominar configuraciones tonales, de acuerdo al número de notas que contiene este elemento. En este pequeño apéndice se detallarán las propiedades de configuraciones tonales que se emplean en la presente tesis (notas individuales, intervalos, acordes y escalas), así como las relaciones que se pueden establecer entre sí para generar una composición musical.

## A.2 Compases y Figuras Rítmicas

Tal como hemos visto, en un pentagrama se pueden expresar melodías, armonías y ritmos. Para expresar altura de sonidos correspondientes a la melodía y armonías, se utilizan las notas. Los ritmos se puede representar mediante lo que se conocen como figuras rítmicas.

Las figuras rítmicas son una representación gráfica de la duración finita de un sonido o de un silencio. La unidad para medir la duración de una nota musical o de un silencio no son los segundos, sino los tiempos. Algunas figuras rítmicas principales se denominan: redonda, blanca, negra, corchea, semicorchea y fusa. Todas las figuras rítmicas tienen una relación binaria entre ellas, de acuerdo a la Figura A.2, donde la negra es la unidad simple.

Las figuras se distribuyen en una composición siguiendo unas divisiones regulares llamadas compases. El compás es la división del tiempo de una composición en partes iguales. El compás está dividido, a su vez, en partes iguales que se denominan tiempos. Sin estas divisiones, sería imposible leer una partitura musical.

Gráficamente, los compases se separan mediante barras. Cuando se termina una composición, se representa mediante una barra doble. Dentro de los compases se pueden colocar todas las figuras que se quieran hasta sumar el número máximo de tiempos de cada compás.

Debido a que en esta tesis no se abordan temas complejos en cuanto a rítmica, el concepto de compás y duración del sonido no se analizará aquí en mayor profundidad.

### **A.3 El concepto de Tono y Semitono y las alteraciones del sonido**

Uno de los conceptos fundamentales de la teoría musical son los intervalos. Los intervalos se basan en unas unidades de medida llamadas tono y semitono. El semitono es la distancia más corta entre dos notas en nuestro sistema musical. En las notas musicales, existe un semitono entre el Mi y el Fa y el Si y el Do. El tono es la distancia equivalente a dos semitonos. Las notas musicales están separadas siempre por un tono salvo los dos casos nombrados anteriormente.

Por diversos motivos, algunos de los cuales se detallarán en el presente documento, la distancia entre las notas puede modificarse utilizando unas grañas denominadas alteraciones, que se colocan al lado izquierdo de las notas musicales que queremos modificar. En realidad, las alteraciones modifican la altura de la nota, lo cual afecta a las distancias entre la nota y el resto. Existen tres tipos principales de alteraciones: sostenido, bemol y becuadro. El sostenido sube un semitono, el bemol baja un semitono y el becuadro anula cualquier alteración que pudiera tener la nota que acompaña. En la práctica, y en algunos casos, al alterar una nota como Mi-Fa o Si-Do, se pasa a la nota siguiente o anterior.

### **A.4 Escalas y Tonalidad**

La escala es una sucesión de notas ordenadas en sentido ascendente o descendente.

Existen varios tipos de escalas (octatónica, que tiene 8 sonidos, cromática, que tiene los 12 sonidos incluidos en la octava, etc.) pero en este trabajo nos concentraremos en un

tipo de escala que es el pilar de la teoría musical occidental: la escala diatónica. Una escala diatónica está formada por siete notas musicales de forma ascendente. Las escalas diatónicas en música occidental tienen dos posibles modos: mayor y menor.

- Escala Mayor: Las escalas diatónicas en Modo Mayor, o Escalas Mayores, tienen un patrón al que imitar. En este caso el patrón es la escala de Do Mayor, que se construye con todas las notas naturales ordenadas de forma ascendente empezando por do. El resto de escalas intentan imitar la misma distribución de tonos y semitonos que la escala de do mayor mediante las alteraciones.
- Escala Menor: Las escalas diatónicas en Modo Menor, o Escalas Menores, tienen también un patrón al que imitar, La menor. Esta escala se construye con todas las notas naturales ordenadas ascendentemente, empezando por la nota La. El resto de escalas intentan imitar la misma distribución de tonos y semitonos que la escala de la menor mediante las alteraciones.

Al intentar imitar la escala de Do mayor o la escala de La menor desde otra nota, es necesario alterar una serie de notas para redistribuir los tonos y semitonos y obtener un modo mayor o menor. Eso produce que siempre se utilicen unas alteraciones fijas al construir una escala de mayor desde la nota La, o una escala menor desde la nota Sol. Este conjunto de alteraciones, ya sean bemoles o sostenidos, se denomina armadura, y se colocan al principio de cada partitura para indicar la escala en la que estamos, ya sea mayor o menor.

Como hemos visto al principio del apéndice, nuestro sistema musical tiene doce sonidos o alturas. Algunos de ellos pertenecen a la escala, y por tanto se denominan notas diatónicas. Otros no aparecerán en la construcción natural de la escala, aunque sí pueden ser utilizados en una composición tonal. Estos sonidos se denominan no diatónicos. En música occidental, estos sonidos no diatónicos suelen tener una función de ornamentación de la composición, o bien de modulación de una tonalidad a otra. Este último concepto no se abordará en este apéndice, pues queda fuera del alcance de este trabajo.

Dependiendo de la escala donde nos encontremos, los sonidos diatónicos cobrarán mayor o menor importancia. Esta organización jerárquica de sonidos siguiendo un determinado criterio se denomina tonalidad. Lo correcto sería decir que la tonalidad es una forma de organizar jerárquicamente las alturas de los sonidos musicales reduciéndolos a una octava, para facilitar el proceso de composición. La escala entonces serían las notas de la tonalidad ordenados en sentido ascendente o descendente con sus respectivas alteraciones en caso necesario, como hemos explicado antes, tal y como se visualiza en la Figura A.3.

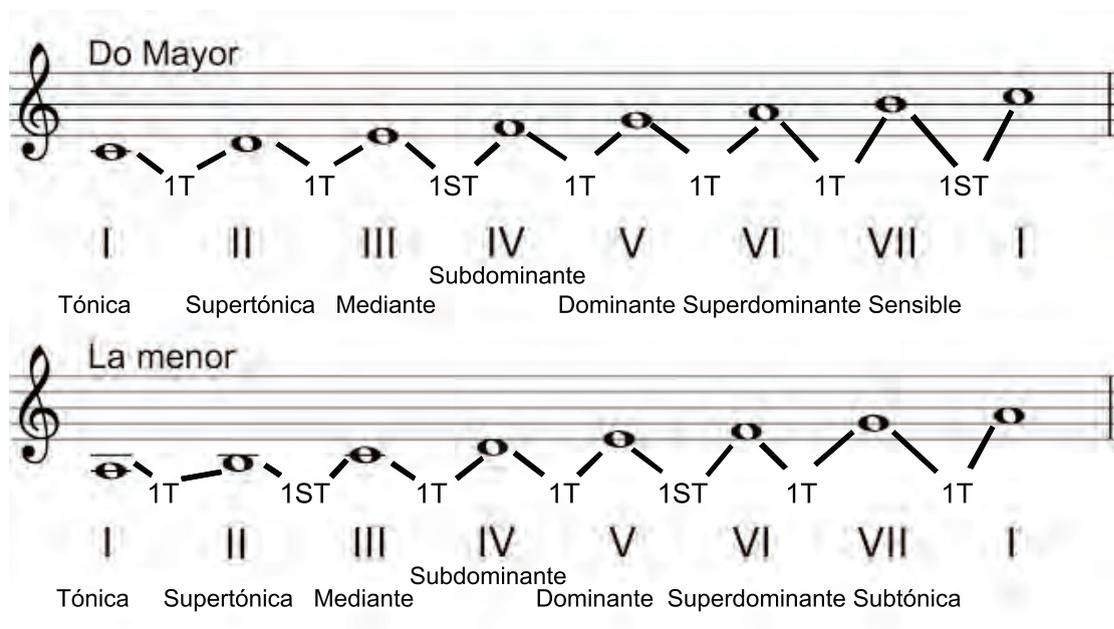


FIGURA A.3: Representación de las escalas modelo mayor y menor, junto con los grados de la escala.

Cada nota de la escala es llamada grado, y cada grado tiene una denominación determinada y se representan mediante números romanos de acuerdo a su situación dentro de la escala (véase Figura A.3). De esta manera, el grado más importante que da nombre a la tonalidad es el grado I, o grado de Tónica. Otros grados importante en una tonalidad son el quinto o el séptimo grado (V o VII) o grados de Dominante y Sensible, ambos provocan la desestabilización de la tonalidad, muy importante para el movimiento melódico y armónico de una composición musical.

Los diferentes grados de la escala y su estructura jerárquica se abordarán en la Sección A.6, donde cobra especial importancia para la denominada progresión de acordes.

## A.5 Intervalos y el concepto de consonancia

La distancia entre dos notas medido en tonos y semitonos se llama intervalo. Un intervalo puede medirse entre dos notas tocadas de forma simultánea (intervalo armónico) o bien entre dos notas consecutivas en una melodía (intervalo melódico) que puede ser ascendente, si la primera nota es más grave que la segunda, o descendente, si la primera nota es más aguda que la segunda.

En cualquier caso, los intervalos también se clasifican dependiendo del número de notas existentes entre una nota y otra, y por su especie, por lo que se pueden clasificar

en: menores (m), Mayores (M), Justos (J), Aumentados (A) y disminuidos (d). Esta clasificación sigue una serie de reglas que veremos a continuación:

- Unísono: Cuando las dos notas consideradas son la misma en la misma octava.
- Intervalos de Segunda: Ocurren cuando hay dos notas entre las notas consideradas. Por ejemplo entre Do-Re o entre Re-Mi. El intervalo de segunda se clasifica como Mayor, si la separación que hay es de un tono (Do-Re) o menor, si la separación es de un semitono (Mi-Fa).
- Intervalos de Tercera: En este caso, hay tres notas entre las notas consideradas. Por ejemplo, entre Sol y Si, inclusive, hay 3 notas (Sol, La, Si), por lo que el intervalo es un intervalo de tercera. Si es un Tono y un semitono es una tercera menor (Re-Fa), si son dos tonos es una Tercera Mayor (Sol-Si).
- Intervalos de Cuarta: En este caso, hay un recorrido de cuatro notas (3 tonos y un semitono) entre las notas consideradas. Los intervalos de Cuarta no son Mayores ni menores, sino Justos.
- Intervalos de Quinta: En este caso, hay un recorrido de cinco notas (4 tonos y un semitono) entre las notas consideradas. Los intervalos de Quinta no son Mayores ni menores, sino Justos.
- Intervalos de Sexta: En este caso, hay seis notas en el recorrido entre las notas consideradas. Si hay dos semitonos entre medias, entonces estamos ante una sexta menor (Mi-Do). Si solo hay un semitono, entonces estamos ante una sexta mayor (Do-La).
- Intervalos de Séptima: En este caso, hay siete notas en el recorrido entre las notas consideradas. Si hay dos semitonos entre medias, entonces estamos ante una séptima menor (Re-Do). Si solo hay un semitono, entonces estamos ante una séptima mayor (Do-Si).
- Intervalos de Octava: En este caso, hay un recorrido de ocho notas (6 Tonos) entre las notas consideradas. Los intervalos de octava son justos.

Cualquier intervalo puede ser también Aumentado o disminuido. Un intervalo puede ser aumentado cuando tiene un semitono más que el que debería tener en caso de ser justo (en los intervalos de cuarta, quinta u octava) o mayor (intervalos de segunda, tercera, sexta y séptima). Un intervalo puede ser disminuido cuando tiene un semitono menos que el que debería tener en caso de ser justo (en los intervalos de cuarta, quinta u octava) o menor (intervalos de segunda, tercera, sexta y séptima).

2 M ↔ 7 m      2 m ↔ 7 m

2 A ↔ 7 d      3 M ↔ 6 m

3ª m ↔ 6ª M      3ª d ↔ 6ª A

4 J ↔ 5 J      4 A (TT) ↔ 5 d (TT)

FIGURA A.4: Representación los tipos de intervalos que se pueden obtener al considerar notas dentro de una misma octava.

La Figura A.4 nos muestra ejemplos de cada clase de intervalo y su nomenclatura musical. Cabe destacar que cada intervalo aquí mostrado tiene un complementario que se obtiene al invertir dicho intervalo. Por ejemplo, cuando invertimos un intervalo de segunda, obtenemos uno de séptima, cuando invertimos uno de tercera, obtenemos uno de sexta, y cuando invertimos una cuarta, obtenemos una quinta. Es curioso también comprobar que al realizar estas inversiones, sus propiedades también pueden invertirse, de manera que el inverso de un intervalo mayor será un intervalo menor, y el inverso de un intervalo aumentado, será un intervalo disminuido. Los intervalos justos no cambian sus propiedades.

Esta característica hace que los intervalos que son complementarios compartan una serie de propiedades sonoras útiles para conocer la consonancia de los mismos.

Cuando suena un determinado intervalo, este puede producir un sonido agradable, también llamado sonido consonante. Un sonido o grupo de sonidos puede considerarse consonante cuando se percibe un sentimiento de unidad o fusión tonal entre notas.

Las relaciones entre frecuencias de dos notas pueden determinar la consonancia de un sonido. Cuando la relación es un número simple, los sonidos son más consonantes que aquellos que no tienen un número simple como valor de relación. Esta teoría fue propuesta por Pitágoras. En particular, esta medida de consonancia ordena los intervalos, como se enumeran a continuación:

1. Unísono y octava.
2. Cuarta y quinta perfecta.
3. Tercera mayor y su complementario, sexta menor.
4. Tercera menor y su complementario, sexta mayor.
5. Segunda menor y séptima mayor.
6. Cuarta aumentada y quinta disminuida, también llamadas tritono (TT).

El concepto pitagórico de intervalos consonantes sigue vigente aún hoy con ciertas modificaciones. Este concepto de consonancia es clave para el movimiento de la música. Este criterio se ha seguido en nuestro trabajo para clasificar las diferentes configuraciones tales como escalas, acordes o notas.

## A.6 Acordes y Funciones Armónicas

Un acorde es un conjunto de tres o más sonidos que (en principio) suenan simultáneamente. El acorde básico está compuesto por tres notas que son: la fundamental, la tercera y la quinta. A este tipo de acorde se le denomina tríada, y son los que se emplean en este trabajo.

Dentro de los acordes tríada, existen cuatro tipos, con algunos ejemplos visualizados en la Figura A.5:

- Acorde Mayor: El acorde mayor está construido al superponer a la primera nota, llamada fundamental, una tercera mayor y una quinta justa (Grado I de la Figura A.5).

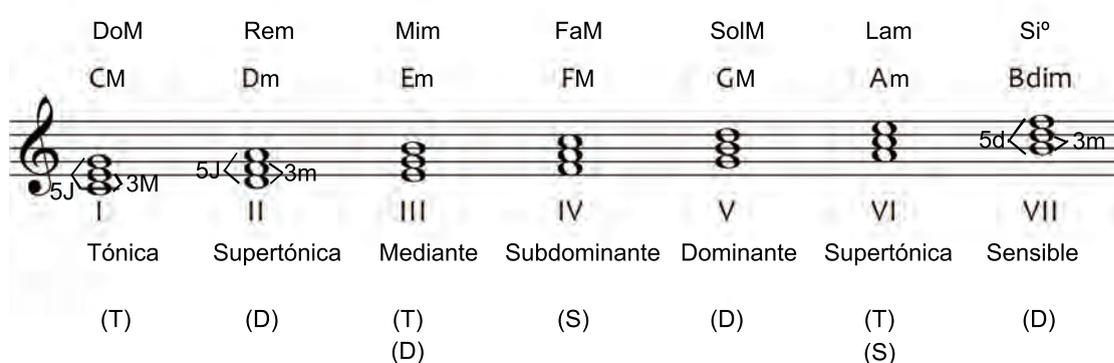


FIGURA A.5: Representación los acordes que se pueden obtener dentro de una tonalidad, en este caso Do/C Mayor. También se visualizan los grados de cada acorde y su función.

- **Acorde Menor:** El acorde menor se construye cuando se superpone a la nota fundamental, una tercera menor y una quinta justa. (Grado II de la Figura A.5)
- **Acorde Aumentado:** Un acorde aumentado está formado al superponer una tercera y una quinta aumentada a la fundamental.
- **Acorde Disminuido:** Un acorde disminuido se forma al superponer una tercera y una quinta disminuida a la fundamental (Grado VII de la Figura A.5).

Dentro de la tonalidad, los acordes también se pueden ordenar siguiendo una jerarquía similar al de las notas individuales. Los acordes también se denominan igual que las notas individuales, mediante su nombre y la numeración romana. De la misma forma que pasaba en los intervalos, los acordes pueden invertirse cambiando de orden las notas. Las inversiones de los acordes no se tratan de forma directa en esta tesis, así que no se ahondará más en este concepto musical.

Los acordes desempeñan una función esencial para el movimiento de la música, cuando estamos hablando de armonía. Este movimiento depende del lugar que ocupan los acordes dentro de la jerarquía. Las funciones armónicas determinan la estructura armónica de una composición musical y están siempre relacionadas con la tonalidad.

Los acordes tríada que se pueden construir en una tonalidad, obedecen a una serie funciones armónicas que se clasifican comúnmente en tres grupos: Tónica (I), Subdominante (IV) y Dominante (V) [Riemann, 1893], que definen juntos la estructura principal de la armonía. En la Figura A.5, los acordes que realizan una función de tónica, están señalados con la letra T, los que hacen función de dominante, con la letra D, y los que hacen función de subdominante con la letra S. Hay que tener en cuenta que

hay determinados acordes que pueden realizar una función u otra, y que esta función dependerá del resto de acordes que están en la progresión.

La tónica o primer grado se considera el centro tonal que conduce a la resolución, la dominante es el siguiente en importancia después de la tónica, mientras que el subdominante puede ser considerado como el dominante inferior.

En la composición musical, los acordes deben ser usados de forma secuencial, formando progresiones que generen tanto un efecto de movimiento como a su vez diferentes matices para acompañar a una determinada melodía. El matiz más importante dentro de la música tonal es la tensión provocada por la sonoridad de la dominante. Esta tensión está dada por el efecto del séptimo grado de la escala que se encuentra dentro del acorde y que provoca una tensión sonora fuerte que obliga a una resolución en la tónica. Este efecto de tensión-resolución es lo que por un lado genera movimiento dentro de la armonía y por otro lado nos indica sobre qué tonalidad está construida la música que se está escuchando. En cambio, el acorde de subdominante representa una función tonal de tensión moderada y no resolutive, ya que en su estructura los acordes que se forman no poseen el efecto tritonal, de esta forma la subdominante complementa y matiza la función tonal de la dominante dentro de la armonía de una pieza musical. En la práctica, las progresiones de acordes tienden a utilizar cíclicamente estas funciones como tónica-subdominante-dominante-tónica y así sucesivamente, para permitir el movimiento musical.

Como hemos podido ver, en una composición musical es esencial identificar la tonalidad principal en la que está compuesta. Una de las formas más sencillas es identificar la armadura de la partitura y los acordes principales para conocer la tónica, la dominante y la subdominante. No obstante, puede ocurrir que solo tengamos acceso a las partituras y tengamos que identificarlo mediante su archivo sonoro. En ese caso, se puede recurrir a la teoría de Krummshal. Esta teoría está basada en el concepto de serie armónica de una nota fundamental. Una serie armónica es una sucesión de los sonidos cuyas frecuencias son múltiplos enteros positivos de la nota fundamental. El primer sonido de la serie, o sonido fundamental, tiene una frecuencia que coincide con la de la nota cuya altura se percibe. El resto de los sonidos se añaden a éste sin alterar su altura aparente, pues el oído funde o integra todos los armónicos en una sola sensación. El segundo sonido de la serie tiene una frecuencia doble de la del primero. Su altura es una octava por encima de aquél. El tercer sonido tiene una frecuencia triple de la del primero, y está en una proporción de 3 a 2 con la del segundo; su altura es una quinta justa por encima de éste. El cuarto sonido tiene una frecuencia doble de la del segundo; su altura será una octava por encima de éste.

La teoría de Krummshal calcula los armónicos de los sonidos que aparecen en una onda sonora, y utilizando la jerarquía de las notas presentada en la tonalidad, establece qué armónicos se repiten y la importancia que tiene cada uno para establecer cuál es la tonalidad y el modo. Esta idea se ha implementado en nuestra tesis para calcular la tonalidad principal del enjambre que aparece en el Capítulo 5.



# Bibliografía

- Aamodt, A. and Plaza, E. (1994). Case-based reasoning: Foundational issues, methodological variations, and system approaches. *AI communications*, 7(1):39–59.
- Abdelhaq, M., Alsaqour, R., and Abdelhaq, S. (2015). Securing mobile ad hoc networks using danger theory-based artificial immune algorithm. *PloS one*, 10(5):e0120715.
- Agmon, E. (1995). Functional harmony revisited: A prototype-theoretic approach. *Music Theory Spectrum*, 17(2):196–214.
- Alberola, J. M., Mulet, L., Such, J., Garcia-Fornes, A., Espinosa, A., and Botti, V. (2007). Operating system aware multiagent platform design. In *Fifth European workshop on multi-agent systems (EUMAS 2007)*, pages 658–667.
- Amiot, E. (2013). The torii of phases. In Yust, J., Wild, J., and Burgoyne, J., editors, *Mathematics and Computation in Music*, volume 7937 of *Lecture Notes in Computer Science*, pages 1–18. Springer Berlin Heidelberg.
- Amiot, E. and Sethares, W. (2011). An algebra for periodic rhythms and scales. *Journal of Mathematics and Music*, 5(3):149–169.
- Anders, T. and Miranda, E. R. (2009). A computational model that generalises schoenberg’s guidelines for favourable chord progressions. In *6th Sound and Music Computing Conference*. Porto, Portugal.
- Anderson, D. T., Zare, A., and Price, S. (2013). Comparing fuzzy, probabilistic, and possibilistic partitions using the earth mover’s distance. *IEEE Transactions on Fuzzy Systems*, 21(4):766–775.
- Argente, E., Botti, V., Carrascosa, C., Giret, A., Julian, V., and Rebollo, M. (2011). An abstract architecture for virtual organizations: The thomas approach. *Knowledge and Information Systems*, 29(2):379–403.
- Argente, E., Botti, V., and Julian, V. (2009). Gormas: An organizational-oriented methodological guideline for open mas. In *International Workshop on Agent-Oriented Software Engineering*, pages 32–47. Springer.

- Atzmueller, M., Becker, M., Kibanov, M., Scholz, C., Doerfel, S., Hotho, A., Macek, B.-E., Mitzlaff, F., Mueller, J., and Stumme, G. (2014). Ubicon and its applications for ubiquitous social computing. *New Review of Hypermedia and Multimedia*, 20(1):53–77.
- Bajo, J., Corchado, J. M., De Paz, Y., De Paz, J. F., Rodríguez, S., Martín, Q., and Abraham, A. (2009). Shomas: Intelligent guidance and suggestions in shopping centres. *Applied Soft Computing*, 9(2):851–862.
- Bajo, J., Corchado, J. M., and Rodríguez, S. (2007). Intelligent guidance and suggestions using case-based planning. In *International Conference on Case-Based Reasoning*, pages 389–403. Springer Berlin Heidelberg.
- Bajo, J., De Paz, J. F., Villarrubia, G., and Corchado, J. M. (2015). Self-organizing architecture for information fusion in distributed sensor networks. *International Journal of Distributed Sensor Networks*, 2015.
- Bell, J. E. and McMullen, P. R. (2004). Ant colony optimization techniques for the vehicle routing problem. *Advanced Engineering Informatics*, 18(1):41–48.
- Bernardes, G., Cocharro, D., Caetano, M., Guedes, C., and Davies, M. E. (2016). A multi-level tonal interval space for modelling pitch relatedness and musical consonance. *Journal of New Music Research*, 45(4):281–294.
- Bianchi, L., Gambardella, L. M., and Dorigo, M. (2002). An ant colony optimization approach to the probabilistic traveling salesman problem. In *International Conference on Parallel Problem Solving from Nature*, pages 883–892. Springer.
- Blackwell, T. (2003). Swarm music: improvised music with multi-swarms. *Artificial Intelligence and the Simulation of Behaviour, University of Wales*.
- Boden, M. (1987). *Artificial Intelligence and Natural Man*. MIT Press, Cambridge.
- Boissier, O. and Gâteau, B. (2007). Normative multi-agent organizations: Modeling, support and control, draft version. *Normative Multi-agent Systems*, (07122).
- Bonabeau, E., Dorigo, M., and Theraulaz, G. (1999). *Swarm intelligence: from natural to artificial systems*. Number 1. Oxford university press, New York City and Oxford.
- Bordini, R., Hübner, J., and Vieira, R. (2005). Jason and the golden fleece of agent-oriented programming. *Multi-agent programming*, pages 3–37.
- Bown, O. (2011). Experiments in modular design for the creative composition of live algorithms. *Computer Music Journal*, 35.
- Bratman, M. (1987). Intention, plans, and practical reason.

- Camarinha-Matos, L. M., Tschammer, V., and Afsarmanesh, H. (2004). On emerging technologies for vo. In *Collaborative Networked Organizations*, pages 207–224. Springer, New York.
- Camazine, S. (2006). Self-organizing systems. *Encyclopedia of cognitive science*.
- Carrascosa, C., Giret, A., Julian, V., Rebollo, M., Argente, E., and Botti, V. (2009). Service oriented mas: an open architecture. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*, pages 1291–1292. International Foundation for Autonomous Agents and Multiagent Systems.
- Chamoso, P., de La Prieta, F., and Villarrubia, G. (2016). Intelligent system to control electric power distribution networks. *Advances on Distributed Computing and Artificial Intelligent Journal (ADCAIJ)*.
- Chang, S.-F., Sikora, T., and Purl, A. (2001). Overview of the mpeg-7 standard. *IEEE Transactions on circuits and systems for video technology*, 11(6):688–695.
- Charron, C., Favier, J., and Li, C. (2006). Social computing. *Forrester Research*.
- Chaudhary, M. D. and Pithadia, P. V. (2014). Multi-feature histogram intersection for efficient content based image retrieval. In *Circuit, Power and Computing Technologies (ICCPCT), 2014 International Conference on*, pages 1366–1371. IEEE.
- Chen, M.-H., Chang, P.-C., and Lin, C.-H. (2014). A self-evolving artificial immune system ii with t-cell and b-cell for permutation flow-shop problem. *Journal of Intelligent Manufacturing*, 25(6):1257–1270.
- Cheng, Q., Lu, X., Liu, Z., and Huang, J. (2015). Mining research trends with anomaly detection models: the case of social computing research. *Scientometrics*, 103(2):453–469.
- Chew, E. (2007). Out of the grid and into the spiral: Geometric interpretations of and comparisons with the spiral-array model. *Computing in Musicology*, 15:51–72.
- Cohen, H. (1995). The further exploits of aaron, painter. *Stanford Humanities Review*, 4(2):141–158.
- Cohn, R. (1997). Neo-riemannian operations, parsimonious trichords, and their “tonnetz” representations. *Journal of Music Theory*, pages 1–66.
- Collins, T., Laney, R., Willis, A., and Garthwaite, P. H. (2016). Developing and evaluating computational models of musical style. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 30(1):16–43.

- Colton, S., Wiggins, G. A., et al. (2012). Computational creativity: The final frontier? In *ECAI*, pages 21–26.
- Cope, D. (1987). An expert system for computer-assisted composition. *Computer Music Journal*, 11(4):30–46.
- Cope, D. (1992). Computer modeling of musical intelligence in emi. *Computer Music Journal*, 16(2):69–83.
- Corchado, J., Bajo, J., and Abraham, A. (2008a). Gerami: Improving the delivery of health care. *IEEE Intelligent Systems*, 23(2):19–25.
- Corchado, J. M., Bajo, J., De Paz, Y., and Tapia, D. I. (2008b). Intelligent environment for monitoring alzheimer patients, agent technology for health care. *Decision Support Systems*, 44(2):382–396.
- Corchado, J. M. and Laza, R. (2003). Constructing deliberative agents with case-based reasoning technology. *International Journal of Intelligent Systems*, 18(12):1227–1241.
- Corchado, J. M., Pavón, J., Corchado, E. S., and Castillo, L. F. (2004). Development of cbr-bdi agents: a tourist guide application. In *Advances in case-based reasoning*, pages 547–559. Springer.
- Corchado, J. M., Villarrubia, G., De Paz, J. F., Rodríguez, S., Zato, C., de la Prieta, F., and Bajo, J. (2013). Practical applications of virtual organizations and agent technology. In *International Conference on Practical Applications of Agents and Multi-Agent Systems*, pages 17–23. Springer.
- Csikszentmihalyi, M. (1997). *Flow and the Psychology of Discovery and Invention*.
- Dartnall, T. (2013). *Artificial intelligence and creativity: An interdisciplinary approach*, volume 17. Springer Science & Business Media.
- David, N., Marietto, M. B., Sichman, J., and Coelho, H. (2004). The structure and logic of interdisciplinary research in agent-based social simulation. *Journal of Artificial Societies and Social Simulation*, 7(3).
- Davidsson, P. (2000). Multi agent based simulation: beyond social simulation. In *Multi-Agent-Based Simulation*, pages 97–107. Springer.
- De Castro, L. N. and Timmis, J. (2002a). An artificial immune network for multimodal function optimization. In *Proceedings of the Congress on Evolutionary Computation, CEC'02.*, volume 1, pages 699–704. IEEE.

- De Castro, L. N. and Timmis, J. (2002b). An artificial immune network for multimodal function optimization. In *Proceedings of the Congress on Evolutionary Computation, CEC'02.*, volume 1, pages 699–704. IEEE, Springer.
- De Castro, L. N. and Von Zuben, F. J. (2000). The clonal selection algorithm with engineering applications. In *Proceedings of GECCO*, volume 2000, pages 36–39.
- De Paz, J. F., Villarrubia, G., Bajo, J., Sirvent, G., and Li, T. (2014). Indoor location system for security guards in subway stations. volume 293, pages 111–119, Salamanca, Spain. Springer Science + Business Media.
- Delgado, M., Fajardo, W., and Molina-Solana, M. (2009). Inmamusys: Intelligent multiagent music system. *Expert Systems with Applications*, 36(3):4574–4580.
- Deng, L. and Tavares, N. J. (2013). From moodle to facebook: Exploring students' motivation and experiences in online communities. *Computers & Education*, 68:167–176.
- Detrain, C. and Deneubourg, J.-L. (2006). Self-organized structures in a superorganism: do ants “behave” like molecules? *Physics of Life Reviews*, 3(3):162–187.
- Dignum, M. (2003). *A model for organizational interaction: based on agents, founded in logic*. Utrecht University Repository.
- Dignum, V. (2009). The role of organization in agent systems. *Handbook of Research on Multi-Agent Systems: Semantics and Dynamics of Organizational Models*, pages 1–16.
- Dignum, V., Meyer, J.-J., Weigand, H., and Dignum, F. (2002). An organization-oriented model for agent societies. In *Proceedings of International Workshop on Regulated Agent-Based Social Systems: Theories and Applications*.
- Dorigo, M., Birattari, M., and Stützle, T. (2006). Ant colony optimization. *IEEE computational intelligence magazine*, 1(4):28–39.
- Dorigo, M. and Stützle, T. (2003). The ant colony optimization metaheuristic: Algorithms, applications, and advances. In *Handbook of metaheuristics*, pages 250–285. Springer.
- Duan, S., Mao, J., Li, J., and Fu, L. (2015). Design implementation and application of swarm intelligence algorithm optimization function simulation platform. In *Software Engineering and Information Technology: Proceedings of the 2015 International Conference on Software Engineering and Information Technology (SEIT2015)*, page 196. World Scientific.

- Dubnov, S. and Assayag, G. (2013). Memex and composer duets: computer-aided composition using style mixing. *Open Music Composers Book*, 2.
- Ebcioğlu, K. (1988). An expert system for harmonizing four-part chorales. *Computer Music Journal*, 12:43–51.
- Ebcioğlu, K. (1990). An expert system for harmonizing chorales in the style of js bach. *The Journal of Logic Programming*, 8(1):145–185.
- Eigenfeldt, A., Bown, O., and Carey, B. (2015). Collaborative composition with creative systems: Reflections on the first musebot ensemble. In *Proceedings of the Sixth International Conference on Computational Creativity June*, page 134.
- Eigenfeldt, A. and Pasquier, P. (2005). Realtime generation of harmonic progressions using controlled markov selection. In *Proceedings of ICC-C-X-Computational Creativity Conference*, pages 16–25.
- Eigenfeldt, A. and Pasquier, P. (2013). Considering vertical and horizontal context in corpus-based generative electronic dance music. In *Proceedings of the Fourth International Conference on Computational Creativity*, page 72.
- Epstein, J. M. and Axtell, R. (1996). *Growing artificial societies: social science from the bottom up*. Brookings Institution Press.
- Erickson, T. and Kellogg, W. A. (2000). Social translucence: an approach to designing systems that support social processes. *ACM transactions on computer-human interaction (TOCHI)*, 7(1):59–83.
- Esparcia, S., Argente, E., Julián, V., and Botti, V. (2014). Gormas: A methodological guideline for organizational-oriented open mas. In *Handbook on Agent-Oriented Design Processes*, pages 173–218. Springer.
- Esteva, M., Rodríguez-Aguilar, J.-A., Sierra, C., García, P., and Arcos, J. L. (2001). On the formal specification of electronic institutions. In *Agent mediated electronic commerce*, pages 126–147. Springer.
- Ferber, J. and Gutknecht, O. (1998). A meta-model for the analysis and design of organizations in multi-agent systems. In *Multi Agent Systems, 1998. Proceedings. International Conference on*, pages 128–135. IEEE.
- Ferber, J., Stratulat, T., and Tranier, J. (2009). Towards an integral approach of organizations in multi-agent systems. *Handbook of Research on MultiAgent Systems: Semantics and Dynamics of Organizational Models*, pages 51–75.

- Ferber, R. (2003). *Information Retrieval: Suchmodelle und Data-Mining-Verfahren für Textsammlungen und das Web*. dpunkt, Heidelberg.
- Ferrara, E., De Meo, P., Catanese, S., and Fiumara, G. (2014). Detecting criminal organizations in mobile phone networks. *Expert Systems with Applications*, 41(13):5733–5750.
- Fukumoto, M. (2014). Creation of music chord progression suited for user’s feelings based on interactive genetic algorithm. In *3rd International Conference on Advanced Applied Informatics (IIAIAAI)*, pages 757–762.
- Galland, S., Gaud, N., Rodriguez, S., and Hilaire, V. (2010). Janus: Another yet general-purpose multiagent platform. In *Seventh AOSE Technical Forum, Paris*.
- Garcá-Pardo, J. A. and Carrascosa, C. (2011). Social welfare for automatic innovation. In *German Conference on Multiagent System Technologies*, pages 29–40. Springer.
- Gaud, N., Galland, S., Hilaire, V., and Koukam, A. (2008). An organizational platform for holonic and multiagent systems. In *Proceedings of Sixth international Workshop on Programming Multi-Agent Systems*.
- Ghasemi, M., Taghizadeh, M., Ghavidel, S., Aghaei, J., and Abbasian, A. (2015). Solving optimal reactive power dispatch problem using a novel teaching–learning–based optimization algorithm. *Engineering Applications of Artificial Intelligence*, 39:100–108.
- Gilbert, N. (2004). *Agent-based social simulation: dealing with complexity*. publisher not identified.
- Girard, C., Ecalle, J., and Magnan, A. (2013). Serious games as new educational tools: how effective are they? a meta-analysis of recent studies. *Journal of Computer Assisted Learning*, 29(3):207–219.
- Giret, A., Julián, V., Rebollo, M., Argente, E., Carrascosa, C., and Botti, V. (2010). An open architecture for service-oriented virtual organizations. In *Programming Multi-Agent Systems*, pages 118–132. Springer.
- Gomez de Silva Garza, A. and Gero, J. (2010). Elementary social interactions and their effects on creativity: A computational simulation.
- Google Inc. DeepDream Project. <https://github.com/google/deepdream>. Online; accessed 9 March 2017.
- Google Inc. Magenta Music Project. <https://github.com/tensorflow/magenta>. Online; accessed 9 March 2017.

- Gratch, J., Mao, W., and Marsella, S. (2006). *Modeling social emotions and social attributions*. Cambridge University Press, Cambridge.
- Gregori, M. E., Cámara, J. P., and Bada, G. A. (2006). A jabber-based multi-agent system platform. In *Proceedings of the fifth international joint conference on Autonomous agents and multiagent systems*, pages 1282–1284. ACM.
- Hare, M. and Deadman, P. (2004). Further towards a taxonomy of agent-based simulation models in environmental management. *Mathematics and computers in simulation*, 64(1):25–40.
- Harte, C., Sandler, M., and Gasser, M. (2006). Detecting harmonic change in musical audio. In *Proceedings of the 1st ACM workshop on Audio and music computing multimedia*, pages 21–26. ACM.
- Herremans, D. and Sörensen, K. (2013). Composing fifth species counterpoint music with a variable neighborhood search algorithm. *Expert systems with applications*, 40(16):6427–6437.
- Herremans, D., Weisser, S., Sörensen, K., and Conklin, D. (2015). Generating structured music for bagana using quality metrics based on markov models. *Expert Systems with Applications*, 42(21):7424–7435.
- Hewlett, W. B. and Selfridge-Field, E. (1997). Midi. In *Beyond MIDI*, pages 41–70. MIT Press.
- Holtzman, S. (1981). Using generative grammars for music composition. *Computer Music Journal*, 5(1):51–64.
- Hoover, A. K., Szerlip, P. A., and Stanley, K. O. (2011). Interactively evolving harmonies through functional scaffolding. In *Proceedings of the 13th annual conference on Genetic and evolutionary computation*, pages 387–394. ACM.
- Horling, B. and Lesser, V. (2004). A survey of multi-agent organizational paradigms. *The Knowledge Engineering Review*, 19(04):281–316.
- Howden, N., Rönnquist, R., Hodgson, A., and Lucas, A. (2001). Jack intelligent agents-summary of an agent infrastructure. In *5th International conference on autonomous agents*.
- Hsu, L.-F., Hsu, C.-C., and Lin, T.-D. (2014). An intelligent artificial system: artificial immune based hybrid genetic algorithm for the vehicle routing problem. *Applied Mathematics & Information Sciences*, 8(3):1191.

- Huang, Z., Zeng, D. D., and Chen, H. (2007). Analyzing consumer-product graphs: Empirical findings and applications in recommender systems. *Management science*, 53(7):1146–1164.
- Hübner, J. F., Boissier, O., Kitio, R., and Ricci, A. (2010). Instrumenting multi-agent organisations with organisational artifacts and agents. *Autonomous Agents and Multi-Agent Systems*, 20(3):369–400.
- Hubner, J. F., Sichman, J. S., and Boissier, O. (2007). Developing organised multiagent systems using the moise+ model: programming issues at the system and agent levels. *International Journal of Agent-Oriented Software Engineering*, 1(3):370–395.
- Huron, D. (1994). Interval-class content in equally tempered pitch-class sets: Common scales exhibit optimum tonal consonance. *Music Perception*, pages 289–305.
- Iglesias, A. and Gálvez, A. (2016). Nature-inspired swarm intelligence for data fitting in reverse engineering: Recent advances and future trends. In *Nature-Inspired Computation in Engineering*, pages 151–175. Springer.
- Jabbarpour, M. R., Malakooti, H., Noor, R. M., Anuar, N. B., and Khamis, N. (2014). Ant colony optimisation for vehicle traffic systems: applications and challenges. *International Journal of Bio-Inspired Computation*, 6(1):32–56.
- Jaiswal, U. and Aggarwal, S. (2011). Ant colony optimization. *International Journal of Scientific & Engineering Research*, 2(7):1–7.
- Jr, I. F., Yang, X.-S., Brest, J., Fister, D., and Fister, I. (2015). Analysis of randomisation methods in swarm intelligence. *International Journal of Bio-Inspired Computation*, 7(1):36–49.
- Julian, V., Carrascosa, C., Rebollo, M., Soler, J., and Botti, V. (2002). Simba: an approach for real-time multi-agent systems. In *Topics in Artificial Intelligence*, pages 282–293. Springer.
- Kennedy, J. (2010). Particle swarm optimization. In *Encyclopedia of Machine Learning*, pages 760–766. Springer.
- Kirke, A. and Miranda, E. (2015). A multi-agent emotional society whose melodies represent its emergent social hierarchy and are generated by agent communications. *Journal of Artificial Societies and Social Simulation*, 18(2):16.
- Krumhansl, C. L. (1990). *Cognitive Foundations of Musical Pitch*. Oxford University Press.

- Krumhansl, C. L. and Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological review*, 89(4):334.
- Lacomme, L., Demazeau, Y., and Dugdale, J. (2010). Clic: an agent-based interactive and autonomous piece of art. In *Advances in Practical Applications of Agents and Multiagent Systems*, pages 25–34. Springer.
- Lester, J. (2002). Rameau and eighteenth-century harmonic theory. *The Cambridge History of Western Music Theory*, pages 753–77.
- Lewin, D. (2001). Special cases of the interval function between pitch-class sets  $x$  and  $y$ . *Journal of Music Theory*, 45(1):1–29.
- Li, X., Mao, W., Zeng, D., and Wang, F.-Y. (2008). Agent-based social simulation and modeling in social computing. In *International Conference on Intelligence and Security Informatics*, pages 401–412. Springer.
- Liu, Y., Wang, K., Guo, H., Lu, Q., and Sun, Y. (2016). Social-aware computing based congestion control in delay tolerant networks. *Mobile Networks and Applications*, pages 1–12.
- López-Ortega, O. and López-Popa, S. I. (2012). Fractals, fuzzy logic and expert systems to assist in the construction of musical pieces. *Expert Systems with Applications*, 39(15):11911–11923.
- Lösch, U., Dugdale, J., and Demazeau, Y. (2009). Requirements for supporting individual human creativity in the design domain. In *Entertainment Computing–ICEC 2009*, pages 210–215. Springer.
- Luck, M., McBurney, P., Shehory, O., and Willmott, S. (2005). *Agent technology: computing as interaction (a roadmap for agent based computing)*. University of Southampton.
- Macedo, L. and Cardoso, A. (2001). Sc-eune-surprise/curiosity-based exploration of uncertain and unknown environments. In *Proceedings of the AISB’01 Symposium on Emotion, Cognition and Affective Computing*, pages 73–81.
- Macedo, L. and Cardoso, A. (2004). Exploration of unknown environments with motivational agents. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pages 328–335. IEEE Computer Society.

- Machado, P., Romero, J., Santos, M. L., Cardoso, A., and Manaris, B. (2004). Adaptive critics for evolutionary artists. In *Applications of evolutionary computing*, pages 437–446. Springer.
- Maher, M. L. (2012). Computational and collective creativity: who’s being creative? In *Proceedings of the 3rd International Conference on Computational Creativity*, pages 67–71. Citeseer.
- Manaris, B., Roos, P., Machado, P., Krehbiel, D., Pellicoro, L., and Romero, J. (2007). A corpus-based hybrid approach to music analysis and composition. In *Proceedings of the National Conference on Artificial Intelligence*, volume 22, page 839. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press.
- Marsden, A. (2000). Music, intelligence and artificiality. In *Music and Artificial Intelligence*, pages 15—28. Miranda ed., Amsterdam: Harwood.
- Martin, A., Jin, C. T., and Bown, O. (2011). A toolkit for designing interactive musical agents. In *Proceedings of the 23rd Australian Computer-Human Interaction Conference*, pages 194–197. ACM.
- McVicar, M., Fukayama, S., and Goto, M. (2014). Autorhythmguitar: Computer-aided composition for rhythm guitar in the tab space. *ICMCSMC (to appear)*.
- Millonas, M. M. (1992). Swarms, phase transitions, and collective intelligence. Technical report, Los Alamos National Lab., NM (United States).
- Mocholi, J. A., Martinez, V., Jaen, J., and Catala, A. (2012). A multicriteria ant colony algorithm for generating music playlists. *Expert Systems with Applications*, 39(3):2270–2278.
- Mojsilovic, A., Hu, H., and Soljanin, E. (2002). Extraction of perceptually important colors and similarity measurement for image matching, retrieval and analysis. *IEEE Transactions on Image Processing*, 11(11):1238–1248.
- Monteith, K., Martinez, T., and Ventura, D. (2012). Automatic generation of melodic accompaniments for lyrics. In *Proceedings of the International Conference on Computational Creativity*, pages 87–94.
- Moroni, A., Manzolli, J., Von Zuben, F., and Gudwin, R. (2000). Vox populi: An interactive evolutionary system for algorithmic music composition. *Leonardo Music Journal*, 10:49–54.
- Morreale, F., De Angeli, A., Masu, R., Rota, P., and Conci, N. (2014). Collaborative creativity: The music room. *Personal and Ubiquitous Computing*, 18(5):1187–1199.

- Mostafa, M. M. and Billor, N. (2009). Recognition of western style musical genres using machine learning techniques. *Expert Systems with Applications*, 36(8):11378–11389.
- Navarro, M., Caetano, M., Bernardes, G., de Castro, L. N., and Corchado, J. M. (2015). Automatic generation of chord progressions with an artificial immune system. In *Evolutionary and Biologically Inspired Music, Sound, Art and Design*, pages 175–186. Springer.
- Pachet, F. (2002). Playing with virtual musicians: The continuator in practice. *IEEE MultiMedia*, 9(3):77–82.
- Pachet, F. (2003). The continuator: Musical interaction with style. *Journal of New Music Research*, 32(3):333–341.
- Paiement, J.-F., Eck, D., and Bengio, S. (2005). A probabilistic model for chord progressions. In *Proceedings of International Conference on Music Information Retrieval*, pages 312–319.
- Pearce, M., Meredith, D., and Wiggins, G. (2002). Motivations and methodologies for automation of the compositional process. *Musicae Scientiae*, 6(2):119–147.
- Pearce, M. T. and Wiggins, G. A. (2007). Evaluating cognitive models of musical composition. In *Proceedings of the 4th international joint workshop on computational creativity*, pages 73–80. Goldsmiths, University of London.
- Pereira, F. C., Machado, P., and Cardoso, A. (1998). Darwinci: Creating bridges to creativity. In *Advances in Artificial Intelligence*, pages 239–248. Springer.
- Pynadath, D. V. and Marsella, S. C. (2005). Psychsim: Modeling theory of mind with decision-theoretic agents. In *IJCAI*, volume 5, pages 1181–1186.
- Quinn, I. (2006). General equal-tempered harmony (introduction and part i). *Perspectives of New Music*, 44(2):pp. 114–158.
- Quinn, I. (2007). General equal-tempered harmony: Parts 2 and 3. *Perspectives of New Music*, 45(1):4–63.
- Ranjbar-Sahraei, B., Shabaninia, F., Nemati, A., and Stan, S.-D. (2012). A novel robust decentralized adaptive fuzzy control for swarm formation of multiagent systems. *Industrial Electronics, IEEE Transactions on*, 59(8):3124–3134.
- Reina, A., Dorigo, M., and Trianni, V. (2014). Collective decision making in distributed systems inspired by honeybees behaviour. In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*, pages 1421–1422. International Foundation for Autonomous Agents and Multiagent Systems.

- Riemann, H. (1893). *Vereinfachte Harmonielehre*. Augener.
- Riesbeck, C. K. and Schank, R. C. (2013). *Inside case-based reasoning*. Psychology Press.
- Roads, C. and Wieneke, P. (1979). Grammars as representations for music. *Computer Music Journal*, 3(1):48–55.
- Robertson, D. and Giunchiglia, F. (2013). Programming the social computer. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 371(1987):20120379.
- Rodríguez, S., De Paz, J. F., Villarrubia, G., Zato, C., Bajo, J., and Corchado, J. M. (2015). Multi-agent information fusion system to manage data from a wsn in a residential home. *Information Fusion*, 23:43–57.
- Rodríguez, S., Julián, V., Bajo, J., Carrascosa, C., Botti, V., and Corchado, J. M. (2011). Agent-based virtual organization architecture. *Engineering Applications of Artificial Intelligence*, 24(5):895–910.
- Rodríguez, S., Zato, C., Corchado, J. M., and Li, T. (2014). Fusion system based on multi-agent systems to merge data from wsn. pages 1–8. IEEE.
- Rubner, Y., Tomasi, C., and Guibas, L. J. (2000). The earth mover’s distance as a metric for image retrieval. *International journal of computer vision*, 40(2):99–121.
- Russell, S. and Norvig, P. (1995). Intelligent agents. *Artificial intelligence: A modern approach*, pages 31–52.
- Saleem, K., Derhab, A., Orgun, M. A., and Al-Muhtadi, J. (2015). Analysis of the scalability and stability of an aco based routing protocol for wireless sensor networks. In *Information Technology-New Generations (ITNG), 2015 12th International Conference on*, pages 234–239. IEEE.
- Sánchez, D. L., Revuelta, J., De la Prieta, F., Gil-González, A. B., and Dang, C. (2016). Twitter user clustering based on their preferences and the louvain algorithm. In *Trends in Practical Applications of Scalable Multi-Agent Systems, the PAAMS Collection*, pages 349–356. Springer.
- Sanz, J. C. (2009). *Lenguaje Del Color: Sinestesia cromática en poesía y arte visual*. H. Blume.
- Schelling, T. C. (2006). *Micromotives and macrobehavior*. WW Norton & Company.
- Schönberg, A. (2006). *The Musical Idea and the Logic, Technique and Art of its Presentation*. Indiana University Press.

- Schuler, D. (1994). Social computing. *Communications of the ACM*, 37(1):28–29.
- Serrano, J. M. and Ossowski, S. (2004). On the impact of agent communication languages on the implementation of agent systems. In *International Workshop on Cooperative Information Agents*, pages 92–106. Springer.
- Shamshirband, S., Anuar, N. B., Kiah, M. L. M., Rohani, V. A., Petković, D., Misra, S., and Khan, A. N. (2014). Co-fais: cooperative fuzzy artificial immune system for detecting intrusion in wireless sensor networks. *Journal of Network and Computer Applications*, 42:102–117.
- Shen, Y., Colloc, J., Jacquet-Andrieu, A., and Lei, K. (2015). Emerging medical informatics with case-based reasoning for aiding clinical decision in multi-agent system. *Journal of biomedical informatics*, 56:307–317.
- Si, M., Marsella, S. C., and Pynadath, D. V. (2006). Thespian: Modeling socially normative behavior in a decision-theoretic framework. In *Intelligent Virtual Agents*, pages 369–382. Springer.
- Sikora, T. (2001). The mpeg-7 visual standard for content description-an overview. *IEEE Transactions on circuits and systems for video technology*, 11(6):696–702.
- Singh, M. P., Rao, A. S., and Georgeff, M. P. (1999). Multiagent systems: A modern approach to distributed artificial intelligence, chapter formal methods in dai: Logic-based representation and reasoning.
- Steedman, M. J. (1984). A generative grammar for jazz chord sequences. *Music Perception*, 2(1):52–77.
- Stiehl, W. D., Lieberman, J., Breazeal, C., Basel, L., Lalla, L., and Wolf, M. (2005). Design of a therapeutic robotic companion for relational, affective touch. In *Robot and Human Interactive Communication, 2005. ROMAN 2005. IEEE International Workshop on*, pages 408–415. IEEE.
- Stock, J. (1993). The application of schenkerian analysis to ethnomusicology: Problems and possibilities. *Music Analysis*, pages 215–240.
- Swartout, W. R., Gratch, J., Hill Jr, R. W., Hovy, E., Marsella, S., Rickel, J., Traum, D., et al. (2006). Toward virtual humans. *AI Magazine*, 27(2):96.
- Thorogood, M., Pasquier, P., and Eigenfeldt, A. (2012). Audio metaphor: Audio information retrieval for soundscape composition. *Proc. of the Sound and Music Computing Cong.(SMC)*.

- Velardo, V. and Vallati, M. (2014). Automatic melody composition and evolution: a cognitive-based approach. *CIM*.
- Videau, S., Bernon, C., Glize, P., and Uribelarrea, J.-L. (2011). Controlling bioprocesses using cooperative self-organizing agents. In *Advances on Practical Applications of Agents and Multiagent Systems*, pages 141–150. Springer.
- Von Ahn, L. (2009). Human computation. In *Design Automation Conference, 2009. DAC'09. 46th ACM/IEEE*, pages 418–419. IEEE.
- Wang, F.-Y. (2007). Toward a paradigm shift in social computing: the acp approach. *Intelligent Systems, IEEE*, 22(5):65–67.
- Wang, F.-Y., Carley, K. M., Zeng, D., and Mao, W. (2007). Social computing: From social informatics to social intelligence. *Intelligent Systems, IEEE*, 22(2):79–83.
- Wooldridge, M. and Jennings, N. R. (1995). Intelligent agents: Theory and practice. *The knowledge engineering review*, 10(02):115–152.
- Wu, J., Peng, D., Li, Z., Zhao, L., and Ling, H. (2015). Network intrusion detection based on a general regression neural network optimized by an improved artificial immune algorithm. *PloS one*, 10(3):e0120976.
- Yang, N.-C., Chang, W.-H., Kuo, C.-M., and Li, T.-H. (2008). A fast mpeg-7 dominant color extraction with new similarity measure for image retrieval. *Journal of Visual Communication and Image Representation*, 19(2):92–105.
- Yang, X.-S. (2010). Firefly algorithm, levy flights and global optimization. In *Research and development in intelligent systems XXVI*, pages 209–218. Springer.
- Yust, J. (2015a). Applications of dft to the theory of twentieth-century harmony. In *Proceedings of the International Conference for Mathematics and Computation in Music*, pages 207–2018.
- Yust, J. (2015b). Schubert's harmonic language and fourier phase space. *Journal of Music Theory*, 59(1):121–181.
- Zaphiris, P. and Ozok, A. A. (2012). Human factors in online communities and social computing. *Handbook of Human Factors and Ergonomics, Fourth Edition*, pages 1237–1249.
- Zato, C., Rodriguez, S., Tapia, D., Corchado, J. M., Bajo, J., et al. (2013). Virtual organizations of agents for monitoring elderly and disabled people in geriatric residences. In *Information Fusion (FUSION), 2013 16th International Conference on*, pages 327–333. IEEE.

- Zaza, T. and Richards, A. (2014). Ant colony optimization for routing and tasking problems for teams of uavs. In *Control (CONTROL), 2014 UKACC International Conference on*, pages 652–655. IEEE.
- Zhang, X., Yi, Z., Yan, Z., Min, G., Wang, W., Elmokashfi, A., Maharjan, S., and Zhang, Y. (2016). Social computing for mobile big data. *Computer*, 49(9):86–90.