



UNIVERSIDAD NACIONAL
AUTÓNOMA DE
MÉXICO

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN CIENCIA E INGENIERÍA DE LA COMPUTACIÓN

**Generación de Melodías Musicales Usando el
Modelo E-R**

T E S I S

QUE PARA OBTENER EL GRADO DE:

MAESTRO EN CIENCIAS (COMPUTACIÓN)

P R E S E N T A:

JUAN MANUEL ALVARADO LÓPEZ

DIRECTOR DE LA TESIS: DR. RAFAEL PÉREZ Y PÉREZ

MÉXICO, D.F.

2009.



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Índice general

Capítulo 1. Introducción	1
1.1. Objetivo	1
1.2. Metas	2
1.3. Motivación, contribución y relevancia	2
1.4. Preguntas de investigación	3
1.5. Organización del contenido	3
Capítulo 2. Estado del Arte	5
2.1. Proceso Creativo:	6
2.1.1. Procesos lógicos y no lógicos	6
2.1.2. La intuición del experto.	9
2.1.3. Creatividad Computacional	11
2.1.4. Creatividad Computacional en Artes.	13
2.2. La Computación Musical	14
2.2.1. Redes Neuronales	16
<i>CONCERT</i>	16
<i>A Connectionist Approach To Algorithmic Composition.</i>	17
2.2.2. Algoritmos Genéticos:	20
<i>A Genetic Algorithm for the Generation of Jazz Melodies</i>	20
<i>Music Through Genetic Algorithms</i>	23
2.2.3. Métodos Simbólicos	25
<i>A Process Algebra for Stochastic Music Composition</i> .	25
Una visión de cómo trabaja Emmy	27
2.3. Música y Emociones	31
2.4. Discusión	32

Capítulo 3. <i>Engagement-Reflection</i>	35
3.1. El proceso creativo para generar historias.	36
3.2. El Modelo Computacional <i>Engagement-Reflection</i>	38
3.2.1. Creación de Estructuras de Conocimiento.	41
3.2.2. El ciclo <i>E-R</i>	41
3.3. Tensión y Emoción	42
3.3.1. Velocidad	49
3.3.2. Diferencia Interválica	50
3.3.3. Variación Interválica	51
3.3.4. La octava promedio	52
3.4. Discusión	53
Capítulo 4. Descripción del Sistema	55
4.1. Introducción	56
4.2. Sistema Difuso	57
4.2.1. Variables Lingüísticas	58
Entradas	59
Salidas	62
4.2.2. Bloques de Reglas	65
Implicación	67
Agregación	67
Defuzzyficación	67
4.3. Construcción de la Base de Conocimiento	67
4.3.1. División de las Melodías	68
4.3.2. Contextos	72
4.3.3. Átomos	74
4.4. Ciclo <i>Engagement-Reflection</i>	76
4.4.1. <i>Engagement</i>	76
4.4.2. <i>Reflection</i>	83
4.5. Resumen de los parámetros del sistema.	88
4.6. Discusión	89
Capítulo 5. Cómo trabaja ERMEG	91
5.1. Construcción de estructuras de conocimiento.	92
5.1.1. Características de las frases	92
5.1.2. Tensión y Emoción	94
5.1.3. Melodía y Ritmo	99

5.1.4.	Contextos y átomos	102
	Contextos	102
	Átomos	102
5.2.	Ciclo <i>engagement-reflection</i>	102
5.2.1.	Inicia Ciclo <i>Engagement</i>	105
	Primer ciclo <i>Engagement</i>	105
	Segundo ciclo <i>Engagement</i>	105
	Tercer ciclo <i>Engagement</i>	108
5.2.2.	Inicia <i>Reflection</i>	111
5.2.3.	Segunda ejecución de <i>engagement</i>	115
	Primer ciclo <i>Engagement</i>	115
	Segundo ciclo <i>Engagement</i>	115
	Tercer ciclo <i>Engagement</i>	116
5.2.4.	Segunda ejecución de <i>reflection</i>	116
5.2.5.	Tercera ejecución de <i>engagement</i>	117
	Primer ciclo <i>Engagement</i>	117
	Segundo ciclo <i>Engagement</i>	118
	Tercer ciclo <i>Engagement</i>	118
5.2.6.	Tercera ejecución de <i>reflection</i>	119
5.3.	Limitaciones	122
5.4.	Discusión	123
Capítulo 6.	Evaluación	125
6.1.	Introducción	125
6.2.	Evaluación de las metas del sistema	125
6.2.1.	Evaluación de la Novedad	126
6.2.2.	Evaluación de la Coherencia	129
6.2.3.	Otras Metas del Sistema	130
6.3.	Retroalimentación	132
6.3.1.	Opiniones	132
6.3.2.	Descripción del Cuestionario	132
6.3.3.	Descripción de los Participantes	134
6.3.4.	Resultados	134
	Primer melodía del sistema	135
	Segunda melodía del sistema	137
	Primer melodía de Metallica	137
	Segunda melodía de Metallica	140

Gráficas Comparativas	142
Última Parte de la Entrevista	142
6.4. Discusión	145
Capítulo 7. Conclusiones	147
Apéndice A. Entrevista	151
Apéndice B. Partituras	159
Bibliografía	165

Índice de figuras

2.1. Pintura realizada por AARON	13
2.2. Pintura realizada por <i>The Painting fool</i>	14
2.3. Juego de dados de Mozart	15
2.4. Esquema de la red de Todd donde se muestran las tres capas de la red y las unidades que las forman.	19
2.5. Etiquetas SPEAC	29
3.1. Esquema general del Modelo Computacional E-R	40
3.2. Mapeo de las características del modelo E-R y las características del sistema ERMEG	43
3.3. Melodía de ejemplo.	45
3.4. Melodía de ejemplo subdividida en frases composicionales. Las líneas verticales indican la división en la melodía.	45
3.5. El sistema selecciona y agrega la Frase Y a la composición en progreso. Después, con la Frase Y incluida en la melodía en progreso, el sistema selecciona y agrega la Frase Z a la melodía.	46
3.6. a) Una frase con muchas notas en una duración de 2 tiempos constituye una frase rápida, el valor obtenido (Velocidad = 4) corresponde a ese rango y b) Una frase con pocas notas en una duración de 8 tiempos constituye una frase lenta, el valor obtenido (0.375) corresponde a ese rango.	47
3.7. a) Una frase que inicia con una nota y que termina con otra $10\frac{1}{2}$ intervalos por encima de la primera tiene una diferencia interválica grande. b) Una frase que inicia con una nota y que termina con otra $3\frac{1}{2}$ intervalos por encima de la primera tiene una diferencia interválica media.	48

3.8. En a) los cambios entre nota y nota son grandes, el promedio de variación es de 8.928 tonos, corresponde con el rango de variación grande; En b) los cambios entre las notas son pequeños, el promedio de variación es de 1.285 tonos, este corresponde con la variación mediana.	49
3.9. Ejemplo de una frase lenta con duraciones de notas de 4, 2 y 2 tiempos. Velocidad = 0.375	50
3.10. Ejemplo de una frase rápida con duraciones de notas de 1/4 de tiempo. Velocidad = 4.0	50
3.11. Ejemplo de una frase con diferencia interválica grande ($10\frac{1}{2}$ tonos).	51
3.12. Ejemplo de dos frases con distinto tipo de variación interválica	52
3.13. Ejemplo de frase de octava promedio alta, la octava promedio es 7.5	52
3.14. Ejemplo de frase de octava promedio baja, la octava promedio es 3	53
3.15. Esquema General de ERMEG	53
4.1. Diseño Conceptual del Sistema Difuso	58
4.2. Funciones de Membresía de “Volumen”	59
4.3. Funciones de Membresía de “Diferencia”	61
4.4. Funciones de Membresía de “Variación”	61
4.5. Funciones de Membresía de “Velocidad”	62
4.6. Funciones de Membresía de “Emoción”	63
4.7. Funciones de Membresía de “Tensión”	64
4.8. ERMEG compone melodías generando secuencias de Frases Composicionales. El sistema agrega la Frase Y y posteriormente la Frase Z a la melodía en progreso.	68
4.9. Divisiones de frases. Las frases uno y dos tienen longitudes de dos tiempos. En las frases tres, cuatro y cinco las duraciones son de cuatro tiempos debido a la presencia de notas que se extienden más allá de los dos tiempos.	70
4.10. Breve melodia	71
4.11. Gráfica de Melodía del ejemplo.	71
4.12. Gráfica de Ritmo del ejemplo	72
4.13. Contexto FC3	73
4.14. Representación de una melodía previa y sus Contextos.	73
4.15. Dos átomos y sus FCs asociadas (o posibles frases a ejecutar).	74

4.16. Generación de contextos y átomos a partir de una secuencia de frases (melodía)	75
4.17. Ciclo <i>Engagement-Reflection</i>	76
4.18. Pasos del ciclo <i>engagement</i>	77
4.19. <i>Match</i> Perfecto entre un átomo y la melodía en progreso.	79
4.20. Equivalencia de valores empleando la constante de resolución	79
4.21. El contexto es equivalente al átomo en 50 %, pues solo la mitad de los puntos comparados son equivalentes	81
4.22. Asociación de los contextos del átomo y de la melodía en progreso usando las constantes de resolución y ACAS	82
4.23. Dos frases que rebasan la máxima distancia permitida. La distancia que separa las notas final de la frase 1 e inicial de la frase 2 es de 13 tonos	85
4.25. <i>Spline</i> que coloca las notas de las frases 1 y 2 y que se dibuja entre ellas en dos compases.	85
4.24. Ruleta para seleccionar las duraciones de las notas en el <i>spline</i>	85
4.26. Dos frases puestas juntas por <i>engagement</i>	86
4.27. Curva del <i>Spline</i> Natural	86
4.28. Modificaciones hechas por <i>reflection</i> . En a) <i>Reflection</i> aplicó una transposición de 12 tonos a la frase 2. En b) Se ha construido un <i>spline</i> natural y se han interpolado notas en dos compases.	87
4.29. Las gráficas se comparan punto a punto para encontrar similitudes de una melodía con otra. En este caso la gráfica de la melodía en progreso es 100 % similar a un fragmento de la melodía previa X. El sistema está copiando.	87
5.1. Ejemplo de una melodía previa.	92
5.2. Melodía previa de ejemplo de la figura 5.1 dividida en frases	92
5.3. Frase 1 de la melodía de muestra	93
5.4. Funciones de membresía de “Velocidad”, ubicación del valor 4 de velocidad, su membresía al conjunto difuso de velocidad alta y su correspondiente valor de membresía.	95
5.5. Funciones de membresía de “Diferencia”, ubicación del valor 2.5 de diferencia. Membresía a los conjuntos difusos diferencia baja y diferencia media y sus correspondientes valores de membresía.	95

5.6. Funciones de membresía de “Variación”, ubicación del valor 1.071 de variación. Membresía a los conjuntos difusos variación baja y variación media y sus correspondientes valores de membresía.	96
5.7. Universo de discurso de las variables Tensión y Emoción y conjuntos difusos mostrados como <i>singletons</i>	98
5.8. Gráfica de Emoción obtenida de las tres frases de la melodía previa de ejemplo.	100
5.9. Gráfica de Tensión obtenida de las tres frases de la melodía previa de ejemplo.	100
5.10. Gráfica de Melodía obtenida de la melodía previa de ejemplo.	101
5.11. Gráfica de Ritmo obtenida de la melodía previa de ejemplo.	101
5.12. Contextos de la melodía previa de ejemplo.	102
5.13. Los átomos de la melodía previa de ejemplo. En ellos podemos encontrar <i>la opción lógica a ser ejecutada</i>	103
5.14. Frase 1 de la melodía de ejemplo que utilizaremos como frase inicial.	104
5.15. La Melodía en progreso después de la primera ejecución de <i>engagement</i>	106
5.16. La Melodía en progreso después de la segunda ejecución de <i>engagement</i> . La pequeña barra en el pentagrama surge como consecuencia de que finalizó el primer compás e indica el principio del segundo.	109
5.17. La Melodía en progreso después de la tercera ejecución de <i>engagement</i>	111
5.18. Modificación hecha por <i>reflection</i> entre la primer y segunda frase de la melodía en progreso. En a) se observa la figura original generada en los ciclos de <i>engagement</i> , en b) se observa la modificación hecha para acercar la segunda frase a la primera.	113
5.19. Frase 2 y 3 de la melodía en progreso con una distancia de separación grande.	113
5.20. Resultado de la segunda modificación hecha por <i>reflection</i>	114
5.21. Resultado de la tercer modificación hecha por <i>reflection</i>	114
5.22. Primer frase agregada en el segundo ciclo <i>engagement</i> . Frase 5	115
5.23. Segunda frase agregada en el segundo ciclo <i>engagement</i> . Frase 6	115
5.24. Tercer frase agregada en el segundo ciclo <i>engagement</i> . Frase 7	116
5.25. Melodía en progreso después de la segunda ejecución de <i>engagement</i>	116

5.26. Melodía después de la segunda ejecución del ciclo E-R completo 117

5.27. Primer frase agregada en el tercer ciclo *engagement*. Frase 8 . 118

5.28. Segunda frase agregada en el tercer ciclo *engagement*. Frase 9 . 118

5.29. Tercer frase agregada en el tercer ciclo *engagement*. Frase 10 . 118

5.30. Melodía en progreso después de la tercera ejecución de *engagement* 119

5.31. Ruleta de probabilidades para seleccionar las duraciones de las notas en el spline. 120

5.32. Curva (*spline*) generada entre las frases 9 y 10 para interpolar notas. 120

5.33. Frases 9 y 10 y las notas interpoladas entre ellas que se ajustan a la curva generada. 121

5.34. Melodía resultante después de tres ciclos E-R completos. . . . 121

6.1. La melodía en progreso se compara sobre todos los niveles de la melodía previa. Cuando existen similitudes entre eventos estos se traslapan. En a) no existe ningún traslape la similitud en esa sección de la melodía previa es 0 %. En b) Existen dos traslapes con lo que la similitud de la pieza en progreso y esa sección es del 50 %. En c) todos los puntos de la melodía en progreso se traslapan con esa sección de la melodía previa, la similitud es del 100 %. En d) nuevamente hay dos traslapes y la similitud es del 50 %. 127

6.2. En a) se muestran las dos frases originales puestas juntas por el sistema. En b) se ha aplicado una transposición a la segunda frase para acercarla a la primera. La distancia que las separa aún rompe la coherencia. En c) se ha aplicado un spline para interpolar notas y la coherencia se logra. 131

6.3. Respuestas para la pregunta “¿La melodía que acaba de escuchar le gustó?” para la composición 1 del sistema. 135

6.4. Respuestas para la pregunta “¿La melodía escuchada fue compuesta por?” para la composición 1 del sistema. 136

6.5. Respuestas para la pregunta “La composición que acaba de escuchar le parece:” para la composición 1 del sistema. 136

6.6. Respuestas para la pregunta “¿La melodía que acaba de escuchar le gustó?” para la composición 2 del sistema. 137

6.7. Respuestas para la pregunta “¿La melodía escuchada fue compuesta por?” para la composición 2 del sistema. 138

6.8. Respuestas para la pregunta “La composición que acaba de escuchar le parece:” para la composición 2 del sistema.	138
6.9. Respuestas para la pregunta “¿La melodía que acaba de escuchar le gustó?” para la composición 1 de Metallica.	139
6.10. Respuestas para la pregunta “¿La melodía escuchada fue compuesta por?” para la composición 1 de Metallica.	139
6.11. Respuestas para la pregunta “La composición que acaba de escuchar le parece:” para la composición 1 de Metallica.	140
6.12. Respuestas para la pregunta “¿La melodía que acaba de escuchar le gustó?” para la composición 2 de Metallica.	141
6.13. Respuestas para la pregunta “¿La melodía escuchada fue compuesta por?” para la composición 2 de Metallica.	141
6.14. Respuestas para la pregunta “La composición que acaba de escuchar le parece:” para la composición 2 de Metallica.	142
6.15. Gráfica comparativa para la pregunta 1 de las cuatro composiciones	143
6.16. Gráfica comparativa para la pregunta 2 de las cuatro composiciones	143
6.17. Gráfica comparativa para la pregunta 3 de las cuatro composiciones	144
B.1. Melodía 1 del ERMEG	160
B.2. Melodia 2 de ERMEG	161
B.3. Melodia 1 de Metallica	162
B.4. Melodía 2 de Metallica	163

Índice de tablas

2.1. Tabla comparativa de los sistemas de composición automática analizados.	32
4.1. Estadísticas del Proyecto	57
4.2. Grados de membresía a los términos bajo, moderado y alto para un nivel de volumen de 25 dB.	59
4.3. Entradas del Sistema difuso	60
4.4. Puntos de definición de las funciones de membresía de “Diferencia”	61
4.5. Puntos de definición de las funciones de membresía de “Variación”	62
4.6. Puntos de definición de las funciones de membresía de “Velocidad”	62
4.7. Salidas del Sistema difuso	63
4.8. Puntos de definición de las funciones de membresía de “Emoción”	64
4.9. Puntos de definición de las funciones de membresía de “Tensión”	64
4.10. Listado de notas y duraciones de la melodía del ejemplo	71
4.11. Resumen de parámetros del sistema.	88
5.1. Valores obtenidos aplicando las reglas de membresía.	97
5.2. Valores de tensión y emoción para las tres frases de la melodía previa del ejemplo	99
5.3. Contexto Inicial	104
5.4. 17 átomos seleccionados por el sistema en la primera ejecución de <i>engagement</i> , pues su porcentaje de similitud con el contexto actual de la melodía está dentro del rango de la constante ACAS.106	

5.5.	37 átomos seleccionados por el sistema en la segunda ejecución de <i>engagement</i> , pues su porcentaje de similitud con el contexto actual de la melodía está dentro del rango de la constante ACAS.	107
5.6.	53 átomos seleccionados por el sistema en la tercera ejecución de <i>engagement</i> , pues su porcentaje de similitud con el contexto actual de la melodía está dentro del rango de la constante ACAS.	109
6.1.	Orden de preferencias promedio	144

Capítulo 1

Introducción

La tecnología computacional ha tenido un enorme avance. La cantidad de aplicaciones que en nuestros días tienen las computadoras han sobrepasado por mucho las ideas de sus creadores y las aplicaciones para la generación de música no podían quedar fuera. Esta es una actividad que ha tenido poca difusión. Los textos destinados al conocimiento de esta tecnología son pocos comparados con los de otras áreas, por ejemplo, redes neuronales, algoritmos genéticos, etc. En esta investigación se aborda el tema de generación de música por computadora y es una contribución al material existente.

1.1. Objetivo

El objetivo de esta investigación es desarrollar un sistema para componer melodías musicales basado en el modelo *Engagement and Reflection* (E-R). El modelo se ha empleado en otras aplicaciones de creatividad computacional y ha tenido resultados notables. Sin embargo, no se ha utilizado para la generación de música. El dominio musical en este trabajo se restringió a melodías monofónicas debido a la complejidad del tema y por lo tanto no hay acompañamientos, ni armonías en las composiciones del sistema. Lo anterior simplifica el proyecto evitando ciertos problemas asociados con la representación de la polifonía. No obstante, las melodías generadas por el programa son realistas e interesantes.

También es objetivo de la investigación representar, en términos computacionales, niveles de tensión y emoción generados en la música. Tales niveles se utilizan en el modelo E-R como elementos de enlace entre frases musicales

para realizar una composición.

Con este proyecto se pretende contribuir al entendimiento de cómo funciona el proceso de composición llevado a cabo sin profundos conocimientos musicales, es decir, el sistema no utiliza durante el proceso de generación de melodías reglas musicales como las de la armonía, contrapunto, etcétera.

1.2. Metas

1. Diseñar e implementar las estructuras y algoritmos necesarios para lograr generar y asociar ideas musicales para componer melodías de acuerdo a las restricciones impuestas por el Modelo E-R
2. Desarrollar una representación en términos computacionales de algunas emociones y tensiones que intervienen en la composición de melodías. En el Modelo E-R no existe ninguna representación computacional de emociones y tensiones para música y es necesario encontrar una adecuada.
3. Que el sistema tome en cuenta el material producido hasta el momento para continuar la composición. Es decir, que todos los eventos pasados en la melodía, que llamaremos contexto, sirvan de guía para generar nuevas ideas musicales. Como se verá en el capítulo 2, la mayoría de los programas para generación de música no incluyen esta característica.
4. Desarrollar rutinas o mecanismos para asegurar que una melodía es coherente. Esto implica definir qué es coherencia y elaborar los algoritmos para cumplirla.
5. Desarrollar rutinas para evaluar la novedad de las melodías generadas por el sistema. Esto implica definir qué es novedad para el sistema.

1.3. Motivación, contribución y relevancia

Es una contribución importante ya que no existe otro sistema para generar melodías utilizando el modelo E-R.

El uso de un sistema de este tipo puede contribuir en diferentes áreas. Por ejemplo el sistema se puede usar en la educación para apoyar el aprendizaje musical y se puede utilizar como instrumento de apoyo a los compositores.

También se puede emplear como herramienta de investigación para entender mejor cómo funciona el proceso creativo. Esto genera nuevo conocimiento por la investigación de las necesidades para generar un nuevo modelo E-R para música.

Otra contribución de esta investigación es el modelado de emociones y tensiones en la música. Este punto es importante, porque la composición musical, así como otras disciplinas artísticas, están basadas en parámetros como las esos, que son difíciles de representar en un sistema de cómputo. De esta manera, con este trabajo se pretende contribuir incorporando al sistema algunas de éstas emociones y tensiones que intervienen en la composición musical.

Para el caso específico de la música esta representación se hará por medio de los parámetros que se pueden encontrar en frases musicales, tales como la velocidad, variación interválica, diferencia interválica, octava en la que se toca la pieza, etc. Tales tensiones y emociones se calcularán por medio de sistemas difusos que a la entrada tendrán los parámetros antes mencionados.

1.4. Preguntas de investigación

- ¿Es posible utilizar el modelo E-R para componer música?
- ¿Es posible desarrollar mecanismos computacionales que puedan generar una representación de las emociones en una melodía?
- ¿Es posible utilizar la representación de emociones para ligar ideas musicales en lugar de teoría musical?
- ¿Es posible generar un programa de computo que represente el proceso de composición como lo realiza alguien que no tiene profundos conocimientos musicales?

1.5. Organización del contenido

El trabajo de tesis está dividido en 7 capítulos que están organizados de la siguiente manera:

- Capítulo 2 Estado del arte. Contiene algunas definiciones dadas por diferentes investigadores sobre lo que es la creatividad. Se muestran

ejemplos de sistemas para la composición automática utilizando diversas técnicas computacionales, como redes neuronales, algoritmos genéticos y métodos simbólicos.

- Capítulo 3 *Engagement-Reflection*. Muestra las características generales del modelo de creatividad computacional *Engagement-Reflection* y establece una relación entre el modelo E-R y el sistema de cómputo producto de esta investigación.
- Capítulo 4 Descripción del sistema. Contiene los detalles del sistema, es decir, las partes que lo conforman y la forma en que interactúan los diferentes módulos para componer una melodía. Se describe el análisis de características musicales por medio de un sistema difuso para obtener representaciones de emociones y tensiones, y luego cómo se utilizan esas emociones y tensiones para componer una nueva melodía ejecutando ciclos de *engagement-reflection*.
- Capítulo 5 Cómo trabaja ERMEG. Muestra ejemplos de ejecución del sistema para completar la descripción del capítulo anterior. En este se muestra la interacción entre los diferentes módulos del sistema.
- Capítulo 6 Evaluación. Se hace una evaluación de la forma de operación de las partes que conforman al sistema y se describen sus resultados. También se evalúa por medio de una encuesta el material producido. En ella varias personas plasmaron sus opiniones acerca de los resultados del sistema. De tal evaluación se desprenden algunas conclusiones.
- Capítulo 7 Conclusiones. Presenta los resultados obtenidos durante el desarrollo de la investigación.

Capítulo 2

Estado del Arte

El objetivo de este capítulo es exponer algunos ejemplos representativos de los trabajos que se han hecho en el área de composición automática de música utilizando diferentes técnicas. Son señaladas algunas opiniones sobre el proceso creativo en general, y en música en particular; se mencionará la importancia de interacción de la música y las emociones. El capítulo está dividido en cinco secciones:

- En la primera parte se exploran puntos de vista sobre el proceso creativo de los humanos y el proceso creativo con computadoras.
- La segunda parte muestra la composición automática basada en redes neuronales. Con este método, las redes reciben como conjuntos de entrenamiento varias melodías y tratan de generalizar las sucesiones de notas. Por ejemplo, tratan de predecir, dado un conjunto de notas pasadas, cual será, con una cierta probabilidad, la siguiente nota en agregarse a la melodía.
- La tercera parte muestra trabajos basados en algoritmos genéticos. Con este método, los individuos de una población de un algoritmo genético al inicio no son composiciones, son conjuntos de sucesiones de notas que evolucionan y se adaptan con una función de evaluación, al final el individuo más “apto” será el elegido. Las funciones de evaluación obedecen a reglas musicales, pero también a los gustos musicales particulares de quienes proponen arquitecturas con estos métodos.
- La cuarta parte muestra ejemplos basados en métodos simbólicos. Con estos métodos, podemos encontrar ejemplos de realizaciones que involu-

cran gramáticas, programación orientada a objetos, en fin sistemas que tratan de manera simbólica el problema de composición automática.

- La quinta parte aborda el tema de música y emociones, la interacción entre ellas y la importancia las emociones en el proceso de composición musical.

2.1. Proceso Creativo:

En el diccionario de la Real Academia de la Lengua Española la definición de creatividad es: 1. f. Facultad de crear. y 2. f. Capacidad de creación.

También se juzga que algo es creativo cuando se produce algo que es nuevo e interesante, aunque por supuesto el mismo producto puede ser evaluado diferente por distintas personas. [Zimmermann, 1999]

La creatividad es un acto de comunicación. Un trabajo de arte, lo es especialmente ya que es producido para ser dado a alguien más, (para mostrar, transmitir, demostrar, etc.). El autor quiere que el receptor esté interesado en ciertas ideas, problemas, circunstancias, dada la naturaleza del ser humano de compartir o intercambiar sus ideas. [Zimmermann, 1999]

2.1.1. Procesos lógicos y no lógicos

Usamos términos como “inspirado”, “intuitivo”, “imaginativo”, “genio”, para explicar habilidades humanas inusuales y mecanismos para ser creativos. Pero, ¿cuáles son los mecanismos importantes para lograr productos que consideramos creativos?

El Dr. Roger Wolcott Sperry, observó el funcionamiento cerebral en pacientes epilépticos que sufrieron comisurotomía total (corte del cuerpo caloso y de la comisura anterior, de forma que se impide la comunicación entre los dos hemisferios cerebrales). Demostró que el hemisferio izquierdo tiene normalmente la dominancia en tareas analíticas y verbales, en tanto que el derecho asume la dominancia en tareas espaciales, en la música y en otros campos. [(Roger Wolcott Sperry, 2008)]¹

Esta idea es compartida por otras personas, hablando de procesos *racionales e intuitivos*, también etiquetados como *analíticos y creativos*. Se dice que

¹Todas las citas de documentos en inglés son traducciones libres del autor de este trabajo.

tales procesos son tan diferentes que son llevados a cabo en diferentes partes del cerebro. [Simon, 1989]

El hemisferio izquierdo (*analítico*) de acuerdo a esto, lleva a cabo el trabajo aburrido, práctico de todos los días; mientras que el hemisferio derecho (*creativo*), es el responsable por los vuelos de imaginación que producen grandiosa música, literatura, ciencia, etc. [Simon, 1989].

Se ha encontrado que los hemisferios del cerebro están especializados para ciertas tareas, pero no se ha encontrado evidencia para decir que así sea para toda tarea.

Herbert Simon [Simon, 1989] menciona que Chester L. Barnard publicó, en el apéndice del libro *Funciones del ejecutivo*, un ensayo titulado *la mente en los asuntos de todos los días*. En él se refiere a los procesos lógicos y no-lógicos. Por procesos lógicos se hace referencia a pensamientos conscientes, los cuales pueden ser expresados en palabras o por otros símbolos, esto es, razonando. Por procesos no-lógicos, se hace referencia a aquellos que no pueden ser expresados en palabras o razonamiento. [Simon, 1989]

Aunque Barnard no provee de un conjunto de criterios formales para distinguir entre la toma de decisiones lógica y no-lógica proporciona una caracterización del fenómeno.

En la toma de decisiones lógica, los objetivos y alternativas son explícitos, las consecuencias de tomar diferentes alternativas están calculadas y estas consecuencias son evaluadas en términos de qué tan cerca están de los objetivos. [Simon, 1989]

En la toma de decisiones no-lógica, la respuesta a una necesidad por una decisión es usualmente rápida, tan rápida como para permitirse un ordenado análisis secuencial de la situación. Quienes toman este tipo de decisiones tienen mucha confianza en lo correctas que son sus decisiones intuitivas y el hecho de tomarlas tan rápidamente es atribuido a su experiencia. Barnard no considera que los procesos no-lógicos de toma de decisiones sean mágicos de ninguna manera. Por el contrario, él creó que están fuertemente relacionados con el conocimiento y experiencia. [Simon, 1989]

El origen de estos procesos no-lógicos yace en factores o condiciones fisiológicas, o en el entorno físico y social impreso en las personas principalmente de forma inconscientemente. También consiste en la masa de hechos, patrones, conceptos, técnicas, abstracciones y generalmente lo que llamamos conocimiento formal o creencias, que son impresas en las mentes de las personas por un esfuerzo más o menos consciente. Este último origen de procesos mentales no-lógicos se incrementa mucho con la experiencia directa, el estudio y

la educación.

Herbert Simon [Simon, 1989] menciona que tales características (conscientes e inconscientes) de los procesos (o toma de decisiones) no-lógicos también forman parte de los procesos lógicos. Ambos procesos incluyen premisas factuales y de valor, y se opera con ellas para formar conclusiones que se convierten en decisiones.

De esta forma si ambos tipos de tomas de decisiones o procesos lógicos y no-lógicos comparten características, y se ha mencionado que existen relaciones entre los tipos de pensamiento intuitivo, creativo y procesos no-lógicos, y entre pensamiento racional, analítico y procesos lógicos, las características son compartidas por ellos también.

Por tanto, la evidencia del pensamiento analítico y creativo, como la de los procesos lógicos y no-lógicos, es esencialmente: la observación de que en todos los asuntos del día, hombres y mujeres hacen juicios, intuyen decisiones razonables y rápidas sin evidencia, y son capaces de reportar el proceso de sus pensamientos que los lleva a sus conclusiones. [Simon, 1989]

Se reporta el proceso incluso si son procesos no-lógicos, pues eso no significa que no tengan explicación ni que no se pueda seguir la secuencia de acciones para la solución de un problema. No se reporta al instante por lo rápido que se presentan las respuestas, pero tales respuestas están basadas en conocimiento y una vasta experiencia, la cual permite esa rápida respuesta que llamamos intuición.

En general el pensamiento analítico y el pensamiento creativo se basan en el conocimiento y la experiencia para la resolución de algún tipo de problema. Entonces, la creatividad no es una facultad especial, es una característica de la inteligencia humana y está cimentada en las capacidades de todos los días, como de asociación de ideas, recordar, percepción, pensamiento analógico, búsqueda de un espacio estructurado del problema y autocrítica reflexiva. Implica no solo a la dimensión cognitiva, si no también a la motivación y emoción. [Boden, 1998]

Como veremos estas ideas se ligan con el modelo Engagement-Reflection, donde Engagement tiene que ver con los procesos no-lógicos para lograr asociar ideas y Reflection tiene que ver con los procesos lógicos en donde se hace una detallada evaluación al material generado.

2.1.2. La intuición del experto.

Newton, a quien se le consideraba una persona “creativa”, fue cuestionado un día de cómo resolvía complicados problemas de ciencia y matemáticas, y dijo: “lidiando constantemente con ellos”. En otra ocasión, atribuía su éxito a “estar parado sobre los hombros de gigantes”. Otro ejemplo es Pasteur que da una variante del primer mecanismo de Newton: “Ideas nuevas vienen a la mente preparada”. [Simon, 2001]

De la revisión de escritos de científicos y del estudio de la psicología cognitiva, que ha generado y probado modelos de procesos creativos, se ha estudiado cómo surgen productos creativos en el laboratorio. (Simon, 2001) Hemos aprendido mucho acerca de los procesos humanos que son usados para resolver problemas, para tomar decisiones, e incluso para crear trabajos de arte y ciencia. Este conocimiento se ha obtenido del estudio de la psicología, de la observación del comportamiento y a través del uso de las computadoras digitales para modelar los procesos de pensamiento, resolución de problemas y toma de decisiones a nivel “Experto”. [Simon, 1989]

Uno de los objetivos de investigación de las ciencias cognitivas ha sido comprender las diferencias entre el comportamiento de expertos y novatos y así posiblemente aprender más acerca de cómo los novatos pueden convertirse en expertos. La finalidad de las investigaciones en inteligencia artificial ha sido construir sistemas de computadora que puedan efectuar tareas profesionales competentemente como puede hacerlo un humano experto.

“Un experto es una persona con mucho potencial para resolver problemas en el dominio de su experticia, pero su potencial no necesariamente se generaliza a otros dominios. La memoria del experto es como una enciclopedia indexada donde las entradas del índice son patrones reconocibles, cada uno de los cuales apunta a un cuerpo de conocimiento relevante y organizado.” [Simon, 2001]. En el ajedrez, por ejemplo, los profesionales deben tener grandes habilidades, conocimiento, memoria e intuición y ellos aplican esto y su gran experiencia para resolver rápidamente una jugada. En esta experiencia el profesional ha acumulado una gran cantidad de patrones que conoce y puede reconocer, y es este reconocimiento el que posibilita al profesional a jugar ajedrez a este nivel. [Simon, 1989] El conocimiento anterior, como se ha dicho, se almacena en patrones y la información asociada con ellos en memoria hace posible esta actuación. Este es el secreto de la intuición de los profesionales del ajedrez (y cualquier profesional). [Simon, 2001] Los expertos parecen dar pasos de gigante en su razonamiento intuitivo comparados con los pasitos de

los novatos. [Simon, 1989]

En un experimento, se mostraba un tablero de ajedrez y luego se retiraban las piezas. Se pidió a una persona (que no supiera mucho de ajedrez) y a un profesional del ajedrez, que volvieran a colocar las piezas en el tablero tal y como se les habían mostrado previamente. El profesional colocó todas las piezas en la misma posición, la otra persona no pudo, y sólo colocó unas cuantas. Se repitió el experimento, pero esta vez las piezas estaban puestas aleatoriamente en el tablero, y nuevamente se pidió que acomodaran las piezas. Ninguno de los dos fue capaz de hacerlo y sólo pudieron colocar unas cuantas, y aproximadamente la misma cantidad de piezas ambos.

En la primera ejecución del experimento, el profesional del ajedrez pudo reconocer en las piezas patrones de un juego, las posibles siguientes jugadas, etc., y por eso fue capaz de reordenarlo completamente. En la segunda ejecución del experimento, al profesional ya no le fue posible encontrar ningún patrón, y se comportó como la otra persona, que no había reconocido patrón alguno. Se ha estimado que un profesional del ajedrez debe ser capaz de reconocer alrededor de 50,000 patrones. En el caso de los universitarios se ha estimado que pueden almacenar en el rango de 50,000 a 200,000 palabras, así que la habilidad para hablar y entender el lenguaje tiene el mismo sabor intuitivo que la habilidad del profesional del ajedrez para jugar rápidamente. [Simon, 1989] Sistemas expertos de computadora que son capaces de igualar el desempeño de un humano en algún limitado dominio pueden ser construidos almacenando en memoria decenas de miles de producciones que son de la forma “*if-then*” y que se evocan de memoria cuando un patrón es reconocido en la situación actual. CADUCEUS y MYCIN son ejemplos de estos sistemas que consisten en un largo número de parejas de producciones “*if-then*”. [Simon, 2001]

La gente creativa ha guardado en memoria una gran cantidad de información acerca de su dominio de creatividad, armada a través de una década o más de esfuerzo, e indexada por patrones que pueden ser reconocidos en las situaciones en las que el proceso creativo se presenta. [Simon, 2001]

Así que, Newton tenía razón, la gente creativa gasta una enorme cantidad de tiempo aprendiendo y pensando en su ciencia o arte. Su creatividad es también incremental, construyéndose sobre la creatividad de aquellos que han estado antes y sobre su propia creatividad. [Simon, 2001]

Como se verá estas ideas también se toman en cuenta en el sistema producto de esta investigación y se incluyen al modelo Engagement-Reflection. El sistema trabajará con patrones que le indicarán la forma de continuar la composición en progreso. Tomando el caso del profesional de ajedrez que pue-

de resolver una jugada rápidamente, el sistema tratará de hacer una “buena jugada musical” en cada momento para lograr una melodía.

2.1.3. Creatividad Computacional

Saunders y Gero [Saunders & Gero, 2002] señalan que la aparente necesidad de definir la naturaleza de la creatividad ha atormentado a la mayoría de los intentos para desarrollar modelos y teorías de los procesos involucrados. La dificultad de esta tarea es clara por el número de definiciones que pueden ser encontradas en la literatura. Algunos investigadores han concluido que tratar de desarrollar una definición simple de creatividad es una tarea infructífera y han buscado maneras para conducir su investigación sin la necesidad de una definición formal.

Aún así Margaret Boden [Boden, 1998] define que una idea creativa es aquella que es novedosa, sorprendente, y valuable (interesante, útil, bella...). Sin embargo “novedoso” tiene dos importantes sentidos. La idea puede ser novedosa con respecto sólo a la mente individual involucrada o, hasta donde tenemos conocimiento de toda la historia previa. La habilidad para producir resultados novedosos de la primera forma puede ser llamada creatividad psicológica (*P-creativity*), y la siguiente, creatividad histórica (*H-creativity*). Pero para los modelos de creatividad con computadoras, y siguiendo las ideas de Boden, nos referiremos a otro tipo de creatividad: la creatividad computacional (*C-creativity*). En este modo de creatividad, que es la adaptación de la creatividad psicológica (*P-creativity*) de Boden a un programa computacional, se podrá decir que se modela un proceso creativo si el programa es capaz de producir conocimiento que no existía previamente en su base de conocimiento, y que es importante para la salida producida. [Pérez y Pérez & Sharples, 2004].

La creatividad artificial es un estudio comparativo de la creatividad tal como es encontrada en las sociedades de humanos contra la creatividad tal como es modelada computacionalmente en sociedades de agentes [Saunders & Gero, 2002].

Ya que ninguna definición parece ofrecer una perspectiva completa de la creatividad, los investigadores Newell, Shaw y Simon [Newell, Shaw & Simon, 1963] desarrollaron una combinación de cuatro criterios tales que siguiéndolos se pueda categorizar una respuesta o solución como creativa:

1. La respuesta es nueva y útil (ya sea para el individuo o la sociedad)
2. La respuesta demanda que las ideas que habían sido previamente aceptadas sean rechazadas.
3. La respuesta resulta de una intensa persistencia y motivación.
4. La respuesta viene de clarificar un problema que era originalmente vago.

Hay que notar como estos criterios tocan muchos de los estereotipos que son típicamente relacionados con la creatividad: novedad y valor (1), transformación y revolución (2), pasión e impulso (3), visión y entendimiento (4). Estos cuatro criterios también combinan elementos de la perspectiva del productor y del producto. El criterio 1 caracteriza las dos cualidades más importantes de un producto creativo, mientras que los criterios 2-4 caracterizan la actitud y las acciones del productor de tal producto. Un producto dado puede satisfacer todos o ninguno de estos criterios, pero esperaríamos que productos que exhiban los cuatro criterios sean ampliamente considerados creativos, mientras que los que exhiban sólo algunos de los criterios serán juzgados con mayor subjetividad y variación. Aunque ningún criterio puede considerarse como necesario o suficiente, el número 1 es tal vez el mayor distintivo de creatividad y sirve de ancla para los demás. Desde la perspectiva computacional el número 1 puede ser considerado como una característica que se debe cumplir, y del 2 al 4 como extras deseables. [“Computational creativity,”2009].

La evaluación de un trabajo creativo es un tema difícil de abordar. De acuerdo a Boden, esto se podría lograr definiendo, en los programas creativos, un espacio conceptual de soluciones aceptadas. [Boden, 1998]. Esto impone el problema de definir tal espacio conceptual adecuadamente. Decir verbalmente qué nos gusta de la música de Bach es muy difícil, pero modelarlo en un programa es incluso más. Y luego decir o modelar lo que no nos gusta puede ser otro tanto complicado [Boden, 1998].

Encontrar una explicación para el fenómeno creativo vía la ciencia de la computación no solo es cuestión de modelar formalismos. Examinar el comportamiento humano, sus metas, limitaciones, apreciaciones, estrategias, y condiciones emocionales, deben preceder a la tarea de modelado. [Zimmermann, 1999]

En general, proveer a un programa con una representación de un espacio conceptual interesante, y con procesos exploratorios apropiados, (para producir resultados creativos) requiere considerable experticia en un dominio



, *Figura 2.1: Pintura realizada por AARON*

específico [Boden, 1998]. Esta idea coincide con la idea de los expertos antes vista.

2.1.4. Creatividad Computacional en Artes.

Un ejemplo de creatividad en las artes es el programa producto de la investigación de Harold Cohen llamado Aarón. Aarón es capaz de crear una pintura completamente independiente cada línea, cada golpe, son determinados por la inspección del estatus de la pintura; sus juicios de desbalance lo llevan a introducir nuevos objetos o modificar los objetos presentes en el dibujo hasta que este alcance un punto en el cuál suelte el lápiz y dé el cuadro por terminado. En la imagen 2.1 se muestra un ejemplo del trabajo de AARON. Los primeros trabajos de Aarón eran orientados a ordenar objetos en el espacio. Hoy en día tiene un gran poder de expresión y puede dibujar siluetas de humanos y los que observan pueden atribuir a estos personajes relaciones sociales. [Simon, 2001]

Otro ejemplo de la creatividad en artes es el trabajo de Simon Colton llamado



, *Figura 2.2: Pintura realizada por The Painting fool*

The painting fool. El sistema es capaz de analizar una imagen digital como una colección de regiones de una pintura y determinar cuáles colores servirán para pintar esas regiones, luego simula el uso de todo tipo de materiales para producir una pintura en un lienzo simulado. Una mejora del sistema implicó que este pueda tomar información acerca de la cara de una persona y escoger los materiales y estilos de pintura apropiados para la emoción mostrada por el o la modelo. Otra mejora hecha es la imaginación del sistema, esto implica la realización de pinturas de cosas que no existen. [Colton, 2009]. En la figura 2.2 se muestra un ejemplo del trabajo de *The painting fool*

Otro caso de creatividad en artes es Emmy [Cope, 1991], un programa de computadora para composición musical desarrollado por David Cope. Este trabajo se describirá con más detalle en la siguiente sección.

2.2. La Computación Musical

Mucho antes de los sistemas que involucran el uso de la computadora para la composición de música existió un sistema mecánico. Wolfgang Amadeus Mozart compuso la obra *Musikalisches Würfelspiel*, singular creación artística en la que el ingenio del músico lo llevó a componer no una pieza para piano

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI
2	96	22	141	41	105	122	11	30	70	121	26	9	112	49	109	14
3	32	6	128	63	146	46	134	81	117	39	126	56	174	18	116	83
4	69	95	158	13	153	55	110	24	66	139	15	132	73	58	145	79
5	40	17	113	85	161	2	159	100	90	176	7	34	67	160	52	170
6	148	74	163	45	80	97	36	107	25	143	64	125	76	136	1	93
7	104	157	27	167	154	68	118	91	138	71	150	29	101	162	23	151
8	152	60	171	53	99	133	21	127	16	155	57	175	43	168	89	172
9	119	84	114	50	140	86	169	94	120	88	48	166	51	115	72	111
10	98	142	42	156	75	129	62	123	65	77	19	82	137	38	149	8
11	3	87	165	61	135	47	147	33	102	4	31	164	144	59	173	78
12	54	130	10	103	28	37	106	5	35	20	108	92	12	124	44	131

, *Figura 2.3: Juego de dados de Mozart*

sino un generador de vales. Esto es, la obra no contiene una partitura para un pequeño vals de 16 compases sino que tiene un sistema que, apoyado en el azar, puede generar un número mucho muy grande de vales diferentes de 16 compases cada uno.

Mozart escribió 176 compases numerados del 1 al 176 y los agrupó en 16 conjuntos de 11 compases cada uno. El procedimiento para generar un vals particular a partir de esta combinación de habilidad en la composición y el uso del azar, consiste en que cada compás del 1 al 16 se selecciona con unos dados, del correspondiente conjunto de 11 compases. Estos 16 conjuntos o columnas de números, que identifican cada uno de los 176 compases se muestran en la figura 2.3

En el encabezado, en la figura 2.3, en romanos aparece el número del compás e identificando cada una de las filas aparece un número entre 2 y 12 que corresponde a la suma de las caras de dos dados que deben ser lanzados para definir

en cada compás, cuál es el elemento que deberá incluirse en la partitura. La obra aparece publicada por primera vez en 1793. [O'Reilly & Ortega, 2008] El uso de las computadoras en la composición musical se introdujo en los años 50. La pieza “*The Illiac Suite for String Quartet*” puede ser considerada como la primera obra compuesta por computadora. Desde aquel momento, las computadoras han jugado un papel clave en diferentes aspectos de la creación musical, desde la síntesis de sonidos complejos, hasta la generación automática. [Alvaro, Miranda & Barros, 2005]

Algunos ejemplos de sistemas representativos para generar música mediante la computadora automáticamente son los siguientes:

2.2.1. Redes Neuronales

CONCERT

Michael Mozer desarrolló *CONCERT* [Mozer, 1994]. *CONCERT* es una red neuronal que compone melodías. El nombre es un acrónimo de *CONnectio-nist Composer of ERudite Tunes*. El conocimiento musical es incorporado en *CONCERT* mediante dos vías. Primero, *CONCERT* es entrenado con un conjunto de melodías de muestra de las cuales se extraen reglas de progresiones de notas y frases, las cuales el autor llamó constricciones melódicas y estilísticas. Segundo, se provee a *CONCERT* con unas bases para juzgar la similitud entre notas, para seleccionar una respuesta (siguiente nota), y para restringir dentro del conjunto de alternativas, las cuales llamó constricciones psicofísicas.

Un modelo completo de composición musical debería describir cada nota con una variedad de propiedades, por ejemplo, tono, duración, acentos, etc., y esto junto a propiedades más globales por ejemplo tempo y cambios dinámicos. En este trabajo el problema fue simplificado, describiendo una melodía como una secuencia de notas.

CONCERT es una red recurrente. Una melodía es presentada a ella una nota a la vez, y su tarea en cada punto en el tiempo es predecir la siguiente nota en la melodía. Durante el entrenamiento, con un conjunto de ejemplos, los pesos de las conexiones en la red son ajustados para que ésta sea capaz de llevar a cabo esta tarea correctamente. La nota actual en la secuencia es representada en la capa de entrada de *CONCERT*, y la predicción de la siguiente nota es presentada en la capa de salida. La salida presentada por la red está afectada por las llamadas constricciones psicofísicas, que están

basadas en análisis psicológicos sobre la percepción de tono que tenemos los humanos.

Se desea que la red haga predicciones acerca de melodías que la gente perciba como agradables. Para esto un ingrediente central de este trabajo ha sido equipar a *CONCERT* con una representación del tono motivada en estudios psicológicos sobre la percepción del sonido. Esto se refiere a que las notas que la gente juzga que son similares deben tener representaciones similares en la red, indicando que la representación en la cabeza concuerda con la representación en la red.

CONCERT en modo de composición puede crear nuevas piezas. Estas piezas, tendrán relación con las melodías que sirvieron como conjunto entrenamiento, y frecuentemente las composiciones contendrán pequeñas secuencias de los ejemplos del conjunto entrenamiento. Esto es porque *CONCERT* ha aprendido los ejemplos del entrenamiento tan bien que en muchos contextos produce una predicción de una nota con probabilidad de 1. Esto significa que el proceso de selección no tuvo la oportunidad de seguir una dirección alterna.

Un serio inconveniente en una representación que está basada en la similitud entre pares de notas presentadas aisladamente, es que quienes las escuchamos no procesamos la notas individuales aisladamente; las notas aparecen en un contexto musical que sugiere una clave musical, y el cual contribuye a la interpretación de las notas. Así pues, la principal deficiencia de *CONCERT*, al momento de componer, es que las composiciones carecen de coherencia global.

A Connectionist Approach To Algorithmic Composition.

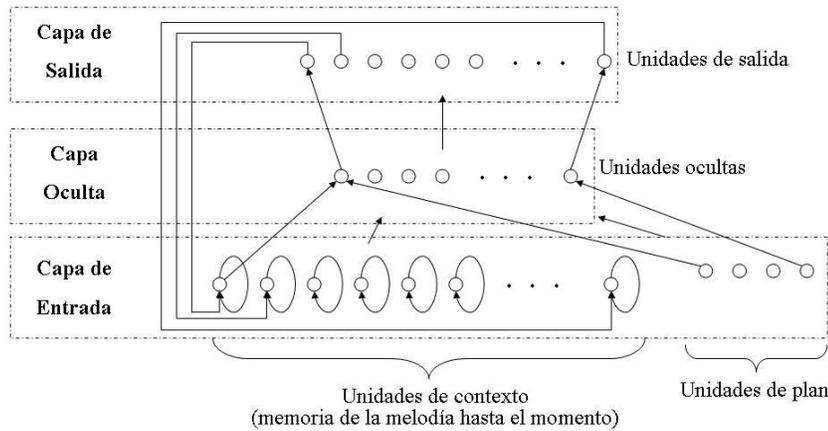
En este trabajo [Todd, 1989] se trata de crear una red que pueda aprender aspectos de la estructura musical y luego se permite a la red utilizar lo que ha aprendido para componer nuevas piezas musicales. Para este trabajo se restringió el dominio musical a la relativamente simple clase de melodías monofónicas.

Todd, eligió una representación de la música usando sólo el tono y la duración de una nota. Las notas en el conjunto de entrenamiento y en las composiciones hechas por esta red, pertenecen sólo a la escala de C. También se omiten otros aspectos dinámicos de la música y de esta forma se simplifica el diseño de la red. Con esta representación de la información musical se inicia el diseño de la red, el cual tiene dos variantes; por una parte se recibe en una ventana

de datos de un tamaño determinado, por ejemplo un compás, un conjunto de datos (notas musicales). De esta forma aprender una melodía puede consistir en asociar cada compás de la melodía con el siguiente. Así como un compás particular se presenta a la entrada de la red, el siguiente compás será producido como salida. Aprender a hacer tales apreciaciones de patrones es algo en lo que las redes son muy eficientes. Más aún, las redes son capaces de generalizar a nuevos patrones que no han aprendido previamente, produciendo salidas razonables en esos casos también. Por otra parte, se tiene el diseño de la red tomando a la música como un fenómeno secuencial, en donde las notas son producidas una detrás de la otra en una sucesión. Con este enfoque, el diseño es una red secuencial que aprende a producir secuencias de notas. Las redes secuenciales aprenden a producir la siguiente nota en una secuencia basándose en una memoria de notas pasadas en la secuencia. Así, se requiere algo de memoria de las notas pasadas en la red secuencial, y esta memoria se provee mediante algún tipo de retroalimentación en la red.

Esta red está formada por tres capas, la primera de ellas es la de entrada, también llamada *capa de contexto*. Esta es llamada así ya que todas las unidades en esta capa almacenan la memoria de las notas pasadas, recibiendo esa información de la capa de salida y retroalimentando en ellas mismas la información de esas notas pasadas. En la capa de entrada además de las unidades de contexto se tienen otras llamadas de plan, éstas contienen información que identifica la melodía que actualmente se está aprendiendo, y en modo de composición identifican la melodía de donde tenemos una respuesta para una entrada dada.

La siguiente es la *capa oculta*. La última es la *capa de salida*. Las unidades de salida determinan lo que la red produce como el siguiente elemento en la secuencia. Cada salida será enviada a la capa de entrada, en donde se agrega y modifica el contexto de la melodía hasta el momento. Estos valores en la red sirven para calcular el siguiente elemento en la secuencia, y el ciclo se repite. En la figura 2.4 se muestra un esquema de la red y sus capas.



, **Figura 2.4:** Esquema de la red de Todd donde se muestran las tres capas de la red y las unidades que las forman.

El número de unidades en las diferentes capas de la red depende de varios factores. El número de unidades de plan debe ser suficiente para especificar diferentes planes para todas las diferentes secuencias que se aprenden. El número de unidades de salida en la red depende de la representación de la secuencia de elementos usada. El número de unidades de contexto depende del tipo de memoria deseada. Finalmente, el número de unidades en la capa oculta depende de lo que la red deba aprender y no sea exactamente especificado. Si hay muy pocas unidades ocultas, la red puede carecer de poder computacional para aprender todo lo necesario. Es una buena idea, no obstante, mantener el número de unidades ocultas tan pequeño como sea posible, porque esto tiende a mejorar la habilidad de la red para generalizar nuevas situaciones (nuevas entradas), lo cual es usualmente importante. Esto ocurre porque la red es forzada a procesar la información a través de un canal estrecho de unidades ocultas, por lo tanto, debe hacer uso únicamente de los aspectos más importantes de la información que procesa. Si tenemos demasiadas unidades ocultas, la red puede usar esas unidades esencialmente para memorizar en lugar de extraer los aspectos importantes que le permiten generalizar desde los datos de entrenamiento.

Todd menciona que de la forma presentada aquí se pueden encontrar promisorios caminos para la exploración de nuevas aproximaciones a la composición algorítmica. Entrenando las redes con melodías seleccionadas y luego usando la habilidad inherente de generalización de las redes, nuevas melodías similares a los ejemplos de entrenamiento originales pueden ser generadas.

La variedad de patrones musicales que pueden ser creados bajo este método indica su potencial de aplicación.

2.2.2. Algoritmos Genéticos:

A Genetic Algorithm for the Generation of Jazz Melodies

George Papadopoulos trabajó en un algoritmo genético para la composición de jazz [Papadopoulos & Wiggins, 1998]. Los algoritmos genéticos han probado ser muy eficientes métodos de búsqueda, especialmente cuando se trabaja con problemas de espacios de búsqueda muy grandes. Esto, junto con su habilidad para proveer múltiples soluciones, lo cual frecuentemente es necesario en dominios de creatividad, los hace una buena elección para hacer la máquina de búsqueda en una aplicación musical.

En una implementación tradicional de algoritmos genéticos, los operadores genéticos no tienen conocimiento del dominio del problema. Ellos generalmente operan a nivel de bit. En contraste, operadores dependientes del problema se diseñan para un problema específico. Por esta razón, pueden lidiar más efectivamente con él. En este trabajo se utilizaron operadores dependientes del problema.

Muchos problemas de algoritmos genéticos se codifican en binario. Una representación simbólica es más fácil de interpretar por el usuario y puede restringir el uso de operadores genéticos a acciones, las cuales tienen más significado en el contexto del dominio de conocimiento. En esta investigación se utilizó una representación simbólica de las estructuras y los datos.

Implementación: La motivación fue encontrar una representación flexible y eficiente para el cromosoma. Se decidió que el cromosoma debía representar los grados en la escala relativos al acorde actual, en lugar de los tonos absolutos. Se usaron la combinación de grados de la escala y su correspondiente acorde, esto significa que no podemos generar notas fuera de la escala, excepto en el caso de que la nota de la melodía se extienda en dos diferentes acordes. El cromosoma entonces es una secuencia de parejas (grado, duración), siendo los silencios distinguidos por la constante *rest* en lugar de grado.

La progresión de acordes, dada por el usuario, en la cual la melodía estará basada, es una secuencia de tripletas de (raíz, tipo de acorde, duración), usando notación musical estándar ejemplo (Bb, maj7, blanca).

En la implementación, la progresión de acordes de entrada determina la du-

ración d de la canción. Una población de tamaño n es un conjunto de n cromosomas representados como parejas (evaluación, genotipo). La preparación de la población consiste de dos pasos: inicialización; luego la evaluación inicial. Para cada cromosoma en la población, una secuencia (genotipo) de m parejas (grado, duración) es generada. Los tonos son escogidos aleatoriamente con probabilidad uniforme. La probabilidad de generar un silencio se estableció a 12.5% el usuario puede especificar las posibles relaciones de las notas y sus respectivas probabilidades. La función de evaluación se usa para calificar a cada genotipo generado aleatoriamente.

Operadores genéticos: Tres clases de mutaciones musicalmente significativas fueron implementadas. Las clases son:

- *Mutaciones locales:* estas mutaciones operan en un fragmento de cromosoma al azar de una longitud también al azar. Se transporta por un número aleatorio de grados; permuta en tiempo; se ordena ascendentemente o descendientemente los tonos; se retrograda; cambian algunos tonos mientras se mantiene el mismo ritmo; varían las duraciones mientras se mantiene el orden de los tonos; y concatena silencios continuos y tonos idénticos. También hay una mutación simple de una nota la cual sólo cambia el tono de una nota hacia arriba o hacia abajo lo cual puede ser de ayuda cuando la melodía necesita sólo pequeños cambios para incrementar mucho la evaluación.
- *Copiar y mutación:* estas mutaciones copian un fragmento aleatoriamente escogido a una posición diferente mientras posiblemente se opera en él con mutaciones locales.
- *Copia restrictiva y mutación:* también escoge fragmentos aleatoriamente pero esta vez de un conjunto dado de posiciones iniciales y con un tamaño constante. La posición inicial puede ser el primer o tercer golpe de cualquier compás, y el tamaño es la mitad de un compás de cuatro cuartos.

Función de evaluación: Se evalúan ocho aspectos para encontrar el valor de la función de evaluación, de los cuales se calcula vía una suma de pesos, la correspondiente adaptación global.

Intervalos grandes: es posible que haya intervalos grandes entre notas consecutivas lo cual puede ser estéticamente irritante; el usuario puede especificar el intervalo más largo permitido entre notas consecutivas. Cada vez que una

nota se encuentra fuera de los límites de intervalo permitidos la suma, de los tamaños de los intervalos que son más grandes que los permitidos, se multiplica por un peso constante de -20.

Asociación de patrones: los escuchas no sólo reconocen sino que también esperan patrones similares en la música. De acuerdo con esto los músicos tienen la tendencia, cuando ellos improvisan, de usar patrones, variarlos, por ejemplo con transposición o cambios de ritmo. En este sistema, el usuario puede especificar si prefiere que el algoritmo trate de encontrar asociaciones exactas o si prefiere una asociación menos exacta de patrones.

Suspensiones: es posible generar suspensiones de notas que van a través de dos acordes consecutivos. Esta parte de la función de evaluación verifica qué pasa en los cambios de acorde. Se consideran cuatro casos, con distintos pesos: si hay una suspensión consonante, significa que la nota es un miembro de ambas escalas (+10); si hay una suspensión disonante, lo que significa que la nota es miembro de la primer escala, pero no de la segunda (-20); no hay suspensión (+ 5); o hay un silencio (+ 5).

Nota en el primer golpe: el primer golpe en un compás es usualmente el más significativo musicalmente. La primer nota puede ser: una nota del acorde (+10); un silencio (+ 10); una nota no de acorde pero si de la escala (-10); o una nota que no es de la escala (-20).

Nota en la mitad del compás: lo mismo que la anterior pero en el tercer *beat*, y por lo tanto menos restringida en sus pesos, excepto en el caso de que la nota no sea de la escala. Sus respectivos pesos son (+5), (+5), (-5), (-20).

Notas largas: el usuario puede especificar lo que considera son notas largas. Las notas largas son principalmente usadas en la música en puntos de éxtasis. De esta forma, es preferible tener notas largas armónicamente estables. Una nota larga puede ser: una nota del acorde (+10); una nota de la escala pero no del acorde (-20); un silencio (-20); una nota consonante suspendida (+20); o una nota disonante suspendida (-20). Los silencios largos son castigados por que dañan la continuidad del sólo.

Curva: esta es una comparación entre la curva del cromosoma y la curva especificada por usuario. El usuario especifica si el promedio de tono de cada compás es menor igual o mayor al anterior compás. Si la curva se cumple, la función de evaluación gana (+30) si no (-30), esto por el número de compases que cumplen o que no cumplen.

Velocidad: similar a la curva, se hace una comparación entre la velocidad de la melodía y la velocidad especificada por el usuario. La velocidad se calcula sumando las notas y silencios en los compases de la melodía y ligándolo a un

valor “lento”, “medio” o “rápido”.

El algoritmo genético converge muy rápidamente a una adaptación alta. Contrariamente, la melodía generada no alcanza un estándar musical aceptable hasta que la versión restrictiva del último tipo de operadores fue implementada y usada. La asociación de patrones en la función de evaluación una vez que fue correctamente ponderada da una sensación de progresión en la melodía.

En conclusión, tal vez la mejor descripción de la función del sistema es que reduce el espacio de búsqueda controlando soluciones que son probablemente no aceptadas musicalmente, por lo tanto incrementa la probabilidad de una salida agradable.

Generating Music Through Genetic Algorithms

Lim Yu-Xi y Lua Kim Teng desarrollaron otro algoritmo genético para generar música [Yu-Xi & Kim Teng, 1996]. Los algoritmos genéticos intentan representar la teoría de la evolución de Darwin, en una serie de pasos computacionales. Los autores han intentado aplicar los algoritmos genéticos a la composición de música, donde las soluciones óptimas son las piezas de mejor sonido.

Estos procesos de computación evolutiva pueden ser aplicados a la solución de problemas. Los autores tratan de aplicar computación evolutiva a la composición de música por computadora. Las presiones evolutivas en este caso fueron nuestro gusto musical, y las piezas que sonaron mejor fueron más probables de propagarse y multiplicarse. El cromosoma es una pieza musical corta, técnicamente conocida como *frase*. Cada nota es un gen, y de esta manera las complejas interacciones entre los genes son ahora fácilmente observables. Ningún gen sonarían bien por sí solo, ya que sólo se podrá determinar calidad en un grupo de genes. La computadora es alimentada con unas pocas reglas para decidir qué tan bien sonó cada cromosoma, y para asignar un valor de adaptación a cada uno. Entonces, unos pocos son escogidos para cruzar y mutación. Luego, éstos continúan y se convierten en la siguiente generación y el ciclo se repite.

Procedimiento: Para empezar se ajustó el alcance a música clásica por su disponibilidad, el número de distintos estilos y su bien documentado conjunto de reglas. Se escogieron piezas de varios autores, y estas piezas musicales fueron analizadas matemáticamente para determinar cuáles combinaciones de tono duraciones y volúmenes eran más frecuentes. Por simplicidad, se

asumió que a una mayor frecuencia de ocurrencias corresponden mejores combinaciones de sonido.

Al principio, la frecuencia de las ocurrencias de las varias combinaciones fue examinada. Por ejemplo, se contó el número de ocurrencias de varios tonos y duraciones para una nota dada. Otro acumulado tomado en consideración fue el atributo presente en una nota y su valor en la siguiente, por ejemplo, la frecuencia de ocurrencias de un tono después de otro tono particular. 80 canciones con alrededor de 10,000 notas fueron analizadas. Este proceso de formación de datos base es uno de los criterios para determinar el valor de adaptación de un cromosoma. Si un cromosoma tiene una secuencia de la cual tiene una alta probabilidad de ocurrencias, entonces tendrá un valor de adaptación mayor.

Varias reglas fueron definidas para restringir el tipo de notas producidas. Estas reglas fueron las siguientes:

- Las notas no pueden estar muy lejos en términos de tono.
- Las notas no pueden ser muy cortas o muy largas.
- Las notas deben ser de la escala de Do mayor.
- No debe haber demasiados silencios.
- Las notas deben ser de ciertos intervalos.
- Notas repetitivas deben ser evitadas.

Cada regla tiene un diferente peso, con lo cuál se controla cuánto afectan al valor final de adaptación.

Inicialmente, los resultados fueron bastante alentadores. Con ligeros cambios a las reglas para determinar la adaptación, el algoritmo generaba piezas musicales razonables que eran algo diferentes de la entrada. Con otras variaciones menores, la música que el algoritmo producía fue esencialmente la misma.

Los autores concluyen que: los algoritmos genéticos pueden ser usados, en un limitado contexto, para generar música. La música producida está fuertemente influenciada por los datos estadísticos de la entrada y no varía significativamente ni aún con el cambio de otros factores. (Yu-Xi, et al 1996)

Otros tipos de música podrían intentarse también, - *pop*, *rock*, *country*, etc. -. El algoritmo fue diseñado para componer sólo frases cortas, pero una modificación podría dar otro algoritmo que pueda manipular y combinar estas frases

para producir una composición completa. Un conjunto de reglas más extenso podría ser adaptado para asegurar mejores resultados musicales. (Yu-Xi, et al 1996)

2.2.3. Métodos Simbólicos

A Process Algebra for Stochastic Music Composition

Este es un trabajo [Ross, 1995] de un modelo llamado *Musical Weighted Synchronous Calculus of Communicating Systems*, o MWSCCS. Los “*Process Algebra*” son formalismos contemporáneos populares para modelar y analizar sistemas concurrentes. Son ampliamente aceptados como efectivos modelos de concurrencia principalmente por el alto grado de abstracción posible con ellos, así como la sensación intuitiva de lenguaje de programación que tienen. En este trabajo se considera la música como compuesta de flujos de comportamientos discretos concurrentes que pueden ser observados en procesos musicales. Estos flujos denotan estructura horizontal de la música - el orden relativo de eventos con respecto a los demás. El trabajo con *process algebra* ha mostrado que este nivel de abstracción es conveniente para analizar complejos comportamientos concurrentes, ya que denota la esencia de la actividad de un sistema sin obstaculizarse con detalles estructurales del sistema en sí. La componente vertical de la música -eventos simultáneos- es naturalmente denotada por el álgebra síncrona. Una importante característica de MWSCCS es su conjunto de primitivas probabilísticas para representar eventos que se adhieren a distribuciones estocásticas. Estas primitivas estocásticas abren la posibilidad para nuevas técnicas en la composición automatizada.

MWSCCS es una mejora del WSCCS de Tofts, el cual transforma en una versión probabilística al SCCS de Milner. MWSCCS le agrega a WSCCS un dominio de acción musical y operadores musicales útiles.

El evento o señal de comunicación más básica e identificable de un proceso es una partícula. En el dominio de la música, un conjunto útil para incluir en el conjunto de partículas, es el conjunto de notas de la escala musical...

Denotamos este conjunto *Notas* $\subseteq A$ como $Notas = \{a_1, a_1^\sharp, a_1^\flat, b_1, b_1^\sharp, b_1^\flat, \dots\}$, donde cada nota tiene el índice de su octava. Cada partícula a tiene un correspondiente complemento \bar{a} .

Informalmente, una partícula con barra representa una salida de comunicación, mientras que la correspondiente sin barra es una entrada de comunicación. Las partículas se usan para construir acciones. Una acción es un evento

que ocurre en un momento de tiempo.

Una característica importante de MWSCCS es la habilidad para asignar frecuencias relativas y prioridades W a los procesos. Por ejemplo, en $2wab + 3wc + 1w\bar{d}$, el término $2wab$ tiene una probabilidad relativa de $\frac{2}{6}$ con respecto a los otros términos, y tiene una prioridad de $w^1 = 1$.

El lenguaje ϵ es definido por la siguiente gramática:

$$\epsilon ::= 0|X|\alpha.E|E\Gamma A|E\setminus A|E[f]|\sum_i w_i E_i|E \times F|\mu_i \tilde{X} \tilde{E}$$

Donde $E, F \in \epsilon$, X es una variable de proceso, $\alpha \in A$, $w_i \in W$, $\alpha \in A \cup W$ y f es una función partícula para renombrar. También se tienen unas reglas de inferencia. Usando estas reglas, transiciones como $E \xrightarrow{\alpha} E'$ son derivables, lo cual denota la transformación de E en E' dada la acción α . Esta α representa un evento compuesto simultáneo, por ejemplo, un acorde.

Los operadores tienen el siguiente significado informal. El proceso nulo 0 es el proceso que no hace nada. El operador prefijo en $\alpha.E$ representa el proceso que puede hacer la acción α , y convertirse en el proceso E . El operador de permiso en $E\Gamma A$ denota el proceso que llevan a cabo las partículas que son miembros del conjunto A , podando todas las acciones que no están en A . El operador de restricción en $E\setminus A$ es el contrario al de permiso. El operador de reetiquetado en $E[f]$ renombra las partículas de acuerdo a f . El operador de elección $\sum_i w_i E_i$ representa la elección de ejecución de un conjunto de procesos E_i . El proceso E_i elegido es influenciado por la frecuencia y prioridad de cualquier expresión ponderada. El operador de composición paralelo en $E \times F$ forza a los procesos concurrentes E y F a comunicarse síncronamente entre ellos.

La definición de procesos y la invocación recursiva también están soportadas.

Operadores más complejos pueden ser derivados con las primitivas vistas.

Lo anterior considera que los eventos musicales sean acciones discretas corriendo síncronamente con un reloj no especificado. El autor señala la posibilidad de usar señales de *Note on* y *Note off* para cada nota, lo cual tiene la ventaja del mapeo a MIDI. Por ejemplo, sea a que representa una Nota on y a' una Nota off. Con esta notación, si suponemos que el reloj global está corriendo a un cuarto de duración de nota entonces la transición:

$$P \xrightarrow{\bar{c}} \xrightarrow{1} \xrightarrow{1} \xrightarrow{1} \xrightarrow{\overline{ec'e'g}} \xrightarrow{1} \xrightarrow{g'} P'$$

Representa una secuencia que consiste de una nota c redonda, una nota e de

un cuarto, y una notas g negra. La acción 1, de espera, se usa para designar intervalos de tiempo. Este proceso se escribe en MWSCCS como:

$$P =_{def} \bar{c}.1.1.1.\overline{ec'}. \overline{e'g}.1.\overline{g'}.P'$$

Como con otras álgebras, construir un sistema musical con MWSCCS se beneficia del comportamiento abstracto observable. Una influencia primaria para su diseño es el uso de un pequeño y poderoso conjunto de operadores para la construcción de procesos.

Una visión de cómo trabaja Emmy

La idea básica detrás de Emmy (Cope, 1991) es lo que Cope ha llamado música recombinante, esto es: la identificación de estructuras recurrentes de varios tipos en las composiciones de los compositores, y la reutilización de estas estructuras en nuevos ordenamientos para así construir una nueva pieza en el “mismo estilo”. Esto puede verse como alimentar a Emmy con las nueve sinfonías de Beethoven, y Emmy dará como resultado la sinfonía número diez de Beethoven. [Hofstadter, 2001]²

El modo de operación de Emmy dado un conjunto de obras de entrada, usualmente todas de un mismo compositor y pertenecientes al mismo tipo (por ejemplo mazurcas), es:

1. Desmenuzar
2. Reensamblar

Esto es lo que Cope ha llamado música recombinante. Pero hay ciertos principios restrictivos que se deben cumplir para garantizar que la salida generada tenga coherencia. Éstos son:

1. Hacer que el flujo local de patrones de cada voz sea similar a aquellos en las obras de entrada.
2. Hacer que el posicionamiento global de los fragmentos sea similar a aquellos en las obras de entrada.

²Emmy es un sistema desarrollado por David Cope. Douglas Hofstadter hizo un análisis de este sistema y la descripción en este apartado toma parte de sus ideas.

La primera de estas constricciones puede estar caracterizada como ensamblaje sintáctico o ensamblaje basado solo en la forma, mientras que la segunda se puede caracterizar como ensamblaje semántico, o ensamblaje basado solo en el contenido. Por separado ninguno de ellos sería muy impresionante pero cuando se usan juntos, son un par muy poderoso.

Ensamblaje sintáctico: La primera es la restricción que involucra a la forma o lo que uno daría en llamar coherencia de flujo, esta restricción se descompone en dos:

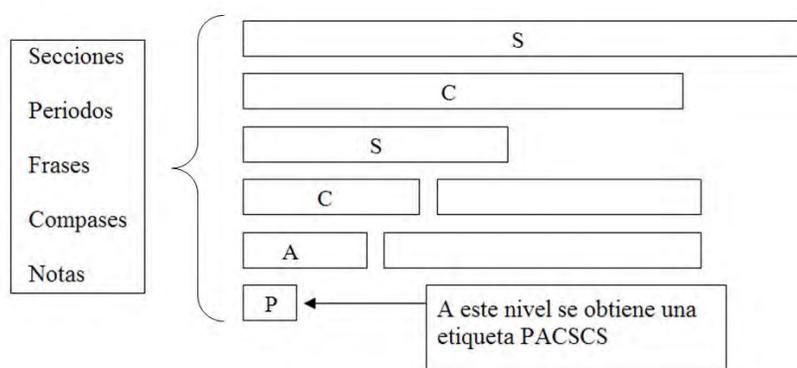
1. Enganchado de voces
2. Combinación de texturas

El enganchado de voces es un requerimiento en el caso de que nosotros tengamos un fragmento f1 y queramos insertar un fragmento f2 después de él. Este funciona así: la nota inicial del fragmento f2 debe coincidir con la siguiente nota melódica del fragmento f1. En otras palabras la transición melódica del fragmento f1 al fragmento f2 debe ser suave. Usualmente si tenemos una cantidad suficiente de obras de entrada podemos tener una alta satisfacción en el enganchado de voces.

La idea básica de la combinación de texturas es que las notas en una secuencia pueden ser movidas arriba o debajo de tono en octavas completas, y se pueden esparcir en el tiempo para semejarse al preexistente patrón local en la pieza que se está componiendo.

Ensamblaje semántico, Lógica de Resolución de Tensión y etiquetas SPEAC: La idea básica es que uno quiere insertar un fragmento en una nueva pieza sólo si la ubicación de la inserción es similar a la ubicación del fragmento donde ocurre en una obra de entrada.

Cope decidió que la ubicación se debe definir de forma que involucre a ambos contextos globales y locales, que de hecho son una serie de contextos anidados empezando desde muy locales (notas, compases) al medio rango (frases) a larga escala (periodos) hasta globales (secciones). A un fragmento en cualquiera de estos niveles (y puede haber cualquier número de niveles estructurales), Cope liga una etiqueta con una de cinco letras S, P, E, A, C, con lo cual intenta capturar lo que Hofstadter ha llamado resolución de estado de tensión de ese fragmento. Las letras usadas para las etiquetas son por las siguientes palabras: *statement*, *preparation* *extensión*, *antecedent* y *consequent*. La asignación de etiquetas procede de lo más local a lo más global, con las etiquetas en las secciones más largas dependiendo únicamente



, *Figura 2.5: Etiquetas SPEAC*

de las etiquetas ya asignadas a las piezas que la forman. En la figura 2.5 se muestra un ejemplo.

De forma que usar estas etiquetas es una muy buena forma de saber en “dónde estás” en una pieza, ya que tensión y resolución en muchos niveles constituyen el verdadero significado de la música.

En nuestra pieza en desarrollo podemos encontrarnos en una ubicación cuya resolución de estado de tensión es PACSCS (de lo más local a lo más global). Las letras PACSCS nos dice en dónde estamos, y así podemos pedir prestado un fragmento de una pieza de entrada, el cual (bajo nuestros criterio de etiquetas) deberá estar localizado en la misma ubicación que dónde lo vamos a insertar, es decir, deberá tener la misma resolución de estado de tensión en la pieza original, es decir PACSCS. En otras palabras, el fragmento que estamos citando está en el “mismo lugar” en la pieza original como en la pieza que estamos componiendo.

Por supuesto también queremos satisfacer la restricción de enganchado de voces sintáctico. Y si en el corpus de entrada tenemos muchos fragmentos que tiene la misma ubicación, eso es bueno pues tenemos muchas opciones por donde seguir, como ya dijimos queremos satisfacer el enganchado de voces, y entonces tenemos que hacer una búsqueda más refinada dentro del grupo de fragmentos que están en la misma ubicación para que también cumplan con el enganchado de voces, si después de refinar la búsqueda nos quedan varios candidatos para ser insertados escogemos uno al azar.

Pero ahora supongamos, que no hay fragmento de las piezas de entrada que tenga la misma resolución de estado de tensión, ¿qué hacer? Cope decidió

que se debía de sacrificar un nivel de tensión y en estas circunstancias lo más sacrificable es lo más global. Así que para el ejemplo citado anteriormente, al no encontrar un fragmento en las mismas condiciones de resolución de estado de tensión, debemos sacrificar el estado más global y nos queda PACSC, ahora volvemos a tratar de encontrar un fragmento adecuado. Si no encontráramos un fragmento adecuado nuevamente se procede de la misma forma (ahora nos quedaría PACS) hasta encontrar un que si nos sirva.

Signatures: Una *signature* (firma) es un patrón de intervalos característico que es recurrente a través de la obra de un compositor, el uso de los cuales nos lleva a componer piezas con un alto nivel de semejanza. Para encontrar las *signatures* Cope pone a Emmy a rastrear en todas las piezas de entrada por pares de secuencias cortas de notas (digamos entre 4 y 12 notas, aunque no hay un límite estricto) cuyos patrones de intervalos corresponden, ya sea, exactamente o aproximadamente. Emmy busca correspondencias exactas, y luego gradualmente relaja la búsqueda, hasta que un número satisfactorio de patrones recurrentes han sido encontrados.

Un punto importante es que estos patrones deben encontrarse entre piezas y no en una misma pieza, por que el compositor en una pieza dada reutiliza sus propios motivos o patrones.

Una vez que las *signatures* han sido encontradas se guardan en una base de datos junto con sus armonías fundamentales, y ya están listas para la inserción como un todo dentro de una nueva pieza. Se podría suponer que la inserción de estos preempaquetados motivos o patrones (pedazos como tal, de una pieza) podrían producir pasajes que sonarían como una auténtica copia de la obra de la cual se extrajo, pero sorprendentemente no, ya que estos “pedazos” de obra son tan genéricos y tan pequeños que aun para los más astutos escuchas pasan sin dar a notar de que obra fueron tomados prestados.

Templagiarism: Esta es una palabra compuesta por las palabras “*template*” y “*plagiarism*” una extensión más abstracta de las *signatures*. En este *templagiarism* lo que hace Emmy es que mientras está buscando en una obra dada de entrada busca si un motivo aparece una o más veces (nuevamente con cierta libertad, el motivo puede ser igual o aproximadamente igual, permitiendo variantes, tales como un desplazamiento tonal, desplazamiento temporal, etc.) Emmy graba los siguientes datos:

1. El desplazamiento tonal de la nueva ocurrencia relativo a la ocurrencia anterior
2. El desplazamiento temporal de la nueva ocurrencia relativo a la ocu-

rrencia anterior

En resumen, Emmy para cada motivo que se repite, graba el patrón de dónde y cuándo, que caracteriza la aparición del motivo. Hay que notar que esto está en un nivel más alto del estilo que un *signature* porque tiene que ver no con el motivo en sí, si no con cómo el motivo es recurrente en una pieza.

2.3. Música y Emociones

En la literatura científica se pueden encontrar trabajos sobre música y emociones. Los investigadores han tratado de averiguar cuáles son las características en la música que nos hacen percibirla de cierto modo. Decir que una pieza musical es alegre, es asignarle una etiqueta, pero ¿qué es lo que nos hace percibirla así? ¿cuáles son las características que tiene una pieza alegre que no tiene una que no lo es? Hay muchos parámetros para juzgar una pieza así. Son tantos, que no es fácil incluirlos todos en un sistema de cómputo. Aún si se hiciera reflejaría solo una verdad parcial, pues lo que para algunos es alegre para otros es triste.

En este trabajo se han tomado en cuenta algunas características musicales para construir un programa que pueda componer melodías basándose en una representación computacional de emociones. Las características extraídas de la música que el sistema utiliza han sido seleccionadas después de revisar la literatura acerca de música y emociones. Se menciona en trabajos científicos por ejemplo: que el tempo influye fuertemente en el tipo de emociones que se perciben en las piezas musicales; tempos rápidos tienden a evocar felicidad y tempos lentos tienden a evocar tristeza [Dalla Bella, Peretz, Rousseau & Gosselin, 2001]. Tempos mayores a 60 bpm (*beats per minute*) evocan tensión y tempos menores evocan suspenso. Tonos altos inducen tensión, un volumen alto causa dolor y uno bajo contribuye a la relajación [Chiu & Kumar, 2003]. La música tiene distintas características (i.e. ritmo y melodía). la variación sistemática de estas características está relacionada con las variaciones de los estados emocionales de los que escuchan. Transformando tales características el contenido emocional puede ser modificado [Oliveira & Cardoso, 2007]. Khalifa et al., [Khalifa, Roy, Rainville, Dalla Bella, & Peretz, 2008] han hecho experimentos en los que encontraron que ni el tempo ni el ritmo solos pueden generar emociones. Lo anterior se logra cuando se combina el ritmo y la melodía.

Juslin y Sloboda, [Juslin & Sloboda, 2001] han mostrado que en la música tonal tonos altos y tempos rápidos generan felicidad; variaciones grandes de tono (saltos) y tonos altos pueden generar ansiedad; tonos graves y tempos lentos pueden generar tristeza.

2.4. Discusión

En este capítulo se han mostrado algunas definiciones de la creatividad en general de los humanos y la implementada en programas de computadora. Se han mostrado algunas condiciones que debería cumplir un sistema cuyos resultados se consideren creativos. También se ha mostrado que es necesaria la adquisición de un gran cuerpo de conocimiento indexado, la adquisición de habilidades de selección y búsquedas heurísticas, y un cuidadoso análisis de nuevos fenómenos revelados por tal búsqueda. [Simon, 2001]

Se han mostrado ejemplos de sistemas de música computacional siguiendo diferentes técnicas. Se analizaron sistemas que trabajan con redes neuronales, algoritmos genéticos y métodos simbólicos. En la tabla 2.1 se muestra una clasificación de los sistemas en la que se describen sus aspectos importantes. Como no todos los sistemas tienen nombre la descripción se hace tomando el nombre del autor.

, **Tabla 2.1:** *Tabla comparativa de los sistemas de composición automática analizados.*

Autor	Descripción
Mozer	El sistema está basado en una red neuronal. El autor ha utilizado una representación musical de características psicofísicas de la música. En modo de composición la predicción que hace la red es la siguiente nota a ser agregada a la melodía. El sistema toma en cuenta el material ya producido para continuar la composición, pero es de sólo una nota, aumentar el contexto significa aumentar el tamaño de la red.
Continúa en la siguiente página	

Tabla 2.1 – Continúa de la página anterior

Autor	Descripción
Todd	Este sistema está basado en una red neuronal. En esta red se selecciona el siguiente evento para continuar la composición con una predicción basada en las relaciones existentes entre partes de la melodía que son aprendidas. Por ejemplo, asociar una nota con la siguiente. La predicción de la red es la siguiente nota de la melodía en progreso. Se considera el material producido para generar el siguiente evento, pero como en el caso anterior, este contexto esta limitado por el tamaño de la red.
Papadopoulos	Este sistema está basado en un algoritmo genético. En este se utilizan operadores dependientes del problema y no se representa en binario a la melodía. La función de evaluación opera con varios criterios musicales que favorecen a la melodía. La salida del sistema es la melodía que mejor se adapte con tal función. Este sistema no toma en cuenta el contexto de la melodía para la composición.
Yu-Xi	Este es otro sistema que utiliza un algoritmo genético. La función de evaluación se basa en criterios que favorecen la melodía, pero no son tan precisos como en el caso anterior. El sistema tampoco toma en cuenta el contexto para la composición.
Ross	Este sistema se basa en un método simbólico. La selección del siguiente evento musical, el cual es una nota, está determinado por el uso de una gramática. El contexto se considera para la generación de nuevos eventos, pero es limitado pues depende de la forma de las reglas en la gramática.
Continúa en la siguiente página	

Tabla 2.1 – Continúa de la página anterior

Autor	Descripción
Cope	También basado en un método simbólico, este sistema produce una melodía nueva utilizando etiquetas SPEAC para conocer la <i>ubicación musical</i> de pequeñas frases de varias melodías. Se extrae de los ejemplos de su base de conocimiento información importante que define el estilo de un compositor particular, con esta información compone en el estilo de ese compositor. Utiliza el contexto para generar un nuevo evento en la melodía, el cual es un conjunto de notas, y se basa en los niveles de tensión y resolución armónica de las melodías.

En los sistemas mencionados hay varias desventajas y ventajas que resaltar. En primer lugar no todos los sistemas toman en cuenta el material producido para la generación del nuevo. Esto es muy importante porque sin esa consideración la música resultante carece de coherencia global o es monótona y repetitiva. Los sistemas que toman en cuenta el contexto de las melodías lo hacen de una manera muy parcial o están limitados para ello. También se debe notar que todos ellos están basados en reglas, o datos estadísticos. Dentro de las ventajas podemos encontrar las representaciones que emplean algunos sistemas, las cuales hacen más fácil la interpretación y manejo de los datos. El uso de etiquetas para conocer la ubicación de una pasaje musical es una ventaja, pues confiere de gran solidez y coherencia a las composiciones. ERMEG se sitúa dentro de los programas que utilizan métodos simbólicos. Se pretende solucionar algunas de las fallas de otros sistemas, por ejemplo no considerar el contexto o estar basados en reglas estrictas. También se ha tomado inspiración de los sistemas descritos en este capítulo para utilizar características que los ayudan a obtener mejores resultados. Por ejemplo, el uso de una representación simple pero descriptiva de la música; utilización etiquetas para conocer la clasificación de frases musicales de una melodía; algunas reglas de los sistemas que sirven para favorecer las melodías resultantes. Además de esto se utilizan algunos resultados de investigaciones en música y emociones para construir una representación de emociones y tensiones de la música para guiar el proceso de composición.

Capítulo 3

Engagement-Reflection

El objetivo de este capítulo es establecer una relación entre el modelo *engagement-reflection* y el programa de computadora producto de este trabajo de investigación. El capítulo consta de las siguientes partes:

- La primera parte trata de las ideas para explicar el proceso creativo de generación de historias propuestas por Mike Sharples.
- La segunda parte describe el modelo computacional *engagement-reflection*, que se basa en las ideas de Sharples y sus principales características. Se menciona cómo se mapean las características antes mencionadas en el programa de computadora producto de este trabajo de investigación.
- La tercera parte describe el modo de obtener una representación computacional de emociones y tensiones para guiar el proceso de composición. Esta representación se obtendrá por medio de características que son medidas en melodías musicales. Se describen las razones para la elección de tales características y lo que representan para el sistema.

3.1. El proceso creativo para generar historias.

Mike Sharples [Sharples, 1999], ha sugerido una serie de ideas para explicar el proceso creativo para generar historias. De acuerdo a lo propuesto por Sharples, el proceso de generación de historias se basa en ciclos entre los estados *engagement* y *reflection*. Durante *engagement* el escritor se concentra y pone toda su atención en generar una cadena de ideas asociadas, que luego convierte a texto. Durante *reflection* el escritor revisa el material generado hasta el momento; lo explora y lo transforma de distintas formas, y genera planes y constricciones que guían la generación de material en sucesivas ejecuciones del ciclo *engagement-reflection*.

En las ideas de Sharples hay cinco aspectos principales que mencionar:

- La importancia de las constricciones: Sharples, hace una distinción entre novedoso y creativo. Por ejemplo: él menciona que para generar oraciones es posible construir un sistema que, siguiendo reglas de gramática, pueda escribir oraciones bien estructuradas que probablemente no habían sido mencionadas. No obstante, el texto generado por este sistema carecería de sentido en la mayoría de los casos. Por tanto, además de ser gramáticamente correcta y novedosa, la escritura debe ser apropiada a la tarea y la audiencia. Sharples, afirma que las constricciones correctas hacen posible la generación de texto apropiado. Esto es lo que distingue a la novedad de la creatividad. Sharples, sugiere que durante *engagement* ellas trabajan como generadores y filtros y por medio de ellas el escritor circunscribe la búsqueda en la memoria a aquellos esquemas mentales apropiados a la tarea.
- Generación y análisis de texto: Sharples, menciona la existencia de dos actividades mentales durante el proceso creativo; una donde el material es generado: *engagement*; y otra donde el material es analizado y evaluado: *reflection*. Durante la primera actividad las características de una idea generada, por ejemplo, personajes, vínculos entre personajes, entornos, etc., motivan la vinculación con una nueva idea cuyas características provocan que una nueva idea más sea agregada, y así sucesivamente, generando una cadena de ideas asociadas. Durante la segunda actividad mental el material es analizado y reestructurado. Este debe satisfacer los objetivos, las creencias y el conocimiento del

escritor acerca del tema.

Gelernter, [Gelernter, 1994] sugiere una idea similar. Gelernter, habla del espectro continuo de una actividad cognitiva que va desde la alta atención (*high focus*), caracterizada por el pensamiento analítico hasta la baja atención (*low focus*), donde las ideas son vinculadas entre sí por medio de las emociones. Gelernter, afirma que la creatividad se reduce al descubrimiento de nuevas analogías (durante *low focus*) que ocurre cuando un pensamiento dispara otro que está relacionado al anterior por emociones compartidas.

Sharpley complementa esta idea sugiriendo que ambos, *low focus* y *high focus*, contribuyen a la creatividad. Cuando se emplean para la producción de textos se convierten respectivamente en los procesos de generación y evaluación mencionados.

- Similitudes entre problemas de diseño y la escritura: Sharpley, señala importantes similitudes que tienen los problemas de diseño y la escritura. Por ejemplo, los problemas de diseño y la escritura no pueden ser completamente especificados y ambos son de duración indeterminada. Esto hace contraste con los problemas clásicos estudiados por la psicología cognitiva y la inteligencia artificial, donde un conjunto definido de objetivos es establecido y el proceso puede ser continuamente evaluado en términos de la proximidad del estado actual hacia el objetivo. Escritores y diseñadores no tienen una función simple para evaluar cada estado. Otro ejemplo es cómo ambos procesos implican tanto buscar como resolver problemas.
- La escritura como diseño creativo, el ciclo *Engagement-Reflection*. Sharpley, establece tres aspectos principales de las ideas que propone:
 - No existe al inicio un objetivo específico o predefinido para desarrollar una historia. Existe un conjunto de constricciones.
 - Mientras la escritura avanza, las constricciones proveen el conocimiento tácito que guía el proceso.
 - El movimiento entre la escritura generativa (*engagement*), guiada por constricciones tácitas, y un estado más deliberativo (*reflection*) forma la máquina cognitiva de la escritura.

Engagement involucra la generación continua de material en un medio externo. Tal material es generado a través de una cadena de ideas asociadas. Durante *reflection* el escritor se detiene para realizar la revisión y/o modificación del material generado para satisfacer los objetivos, las creencias y el conocimiento del escritor. Sharples, sugiere que la interacción entre los estados *engagement* y *reflection* hace avanzar el proceso de escritura de una historia.

3.2. El Modelo Computacional *Engagement-Reflection*

Basándose en estas ideas se desarrolló un modelo computacional y se utilizó en un programa de computadora para la generación de historias llamado MEXICA [Pérez y Pérez & Sharples, 2001]. Las principales características del modelo computacional *Engagement-Reflection* son:

- El usuario provee al sistema de un conjunto de ejemplos que son empleados para construir las estructuras de conocimiento.
- El modelo incluye dos procesos principales: un proceso de generación (conocido como *engagement*) y un proceso de evaluación (conocido como *reflection*). Durante *engagement* el sistema genera material guiado por constricciones; durante *reflection* el sistema evalúa el material generado hasta el momento -lo modifica si es necesario- y como resultado de la evaluación ajusta las constricciones que guían la generación de material durante *engagement*. La salida del sistema es el resultado de la interacción entre los estados *engagement* y *reflection*.
- Una característica importante del modelo es que, siguiendo a Gelernter [Gelernter, 1994], una representación de las emociones es el pegamento que une a las ideas durante *engagement*. De esta forma, en MEXICA, los vínculos emocionales y tensiones entre personajes guían la generación de material durante el desarrollo de la trama [Pérez y Pérez, 2007].
- El modelo provee parámetros para controlar el comportamiento de los diferentes procesos que conforman el ciclo E-R.

En la figura 3.1 se muestra un esquema de las partes que forman el modelo computacional *Engagement-Reflection* (E-R). En la parte superior tenemos

las posibles acciones a ser ejecutadas por un personaje que provee el usuario y los ejemplos previos para que el sistema pueda construir su base de conocimiento. De los ejemplos previos, el sistema extrae diferentes tipos de información, por ejemplo, en que contexto se ha de ejecutar una acción. La acción que el usuario provee es la semilla con la que inicia la ejecución del sistema. De ella también se extrae información para iniciar el proceso de generación. La base de conocimiento contiene información que se utiliza para el desarrollo de material nuevo. El ciclo E-R es el encargado de: la creación de las cadenas de ideas y de la revisión y ajuste de tales ideas en los estados *engagement* y *reflection* respectivamente. Al final del ciclo obtendremos una salida resultado de la interacción entre los estados *engagement* y *reflection*. El sistema MEXICA emplea un tipo de constricciones muy importante: los vínculos emocionales entre los personajes de una historia. Éstos vínculos influyen en las posibles acciones a ejecutar en una historia. Por ejemplo:

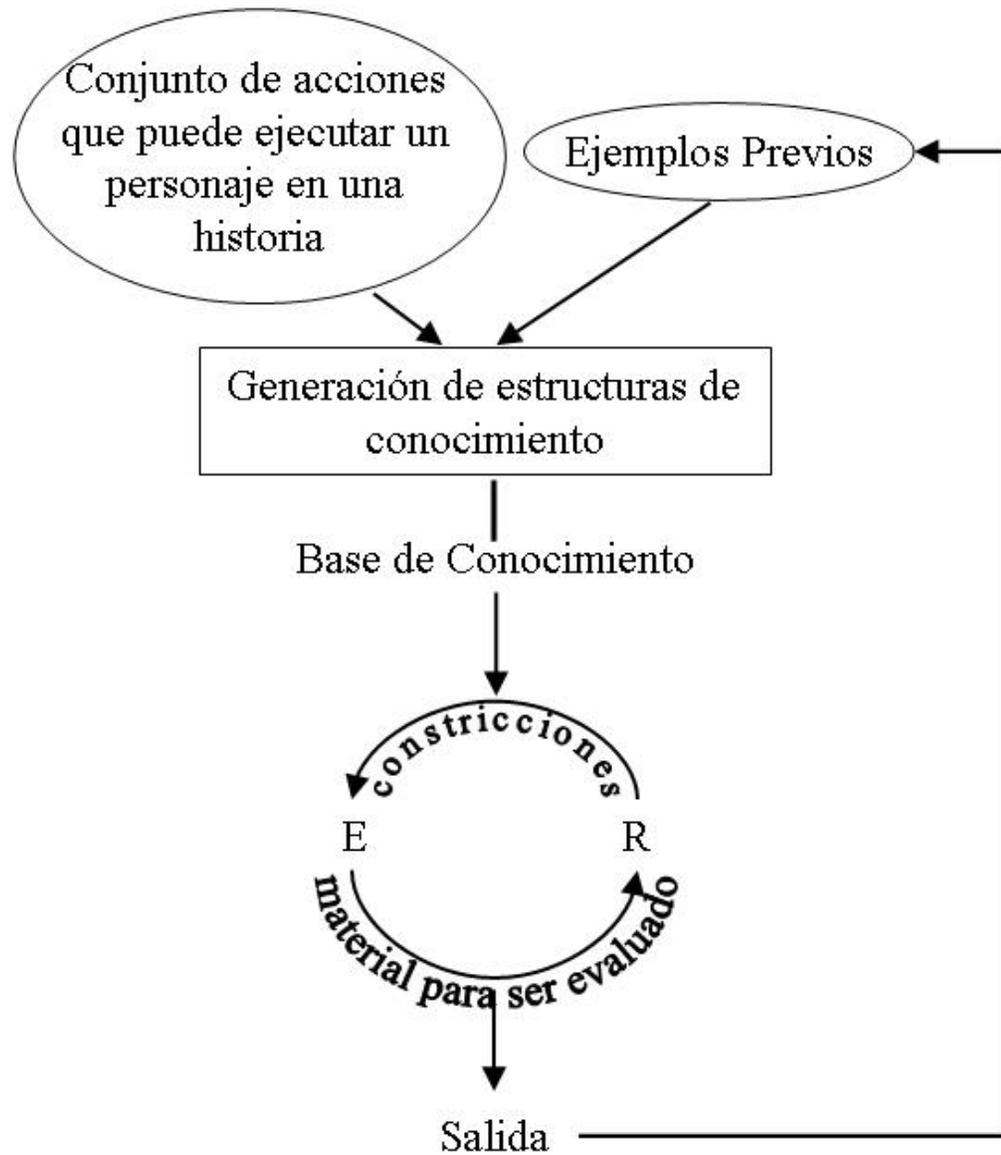
- Si un personaje A ama a un personaje B, puede ser adecuado tener como idea siguiente a este hecho, que el personaje A proteja al personaje B ...
- Podría ser inadecuado que el personaje A lastime al personaje B ...
- más aún, si B lastima a A (accidentalmente o por casualidad), sería entonces lógico, que A deje de amar a B, y entonces la acción de que el personaje A lastime al personaje B tendría sentido.

Durante *reflection* se completa la coherencia entre ideas o acciones que no han quedado completamente satisfechas. Por ejemplo, teniendo las ideas siguientes:

1. El personaje A ama al personaje B ...
2. El personaje A lastima al personaje B ...

Reflection evalúa las acciones y encuentra una inconsistencia, pues no es lógico que esto suceda. Así que, *reflection* inserta una acción entre las dos anteriores para mantener la coherencia de la historia.

1. El personaje A ama al personaje B ...
2. *El personaje B golpeó al personaje A ...*



, *Figura 3.1: Esquema general del Modelo Computacional E-R*

3. El personaje A lastima al personaje B . . .

En los casos anteriores los ejemplos son presentados con fines ilustrativos. Los vínculos emocionales y las acciones pueden ser más complejos. También puede haber más constricciones que sólo los vínculos emocionales entre personajes. Pero, en esencia, esta es la forma en que se utilizan este tipo de constricciones para generar un historia empleando el modelo *engagement-reflection*.

A continuación se describe la generación de estructuras de conocimiento en el modelo E-R.

3.2.1. Creación de Estructuras de Conocimiento.

El modelo *engagement-reflection* requiere de un conjunto de estructuras de conocimiento para funcionar. Esta base de conocimiento se construye utilizando un conjunto de ejemplos llamados ejemplos previos. Las estructuras de conocimiento se construyen de los ejemplos previos que el usuario provee al sistema y corresponden a la experiencia, conocimiento o repertorio del sistema. En un sistema que genera historias, como es el caso de MEXICA, los ejemplos previos son historias. En el caso de la generación de melodías musicales los ejemplos previos son melodías. Las estructuras de conocimiento contendrán información relevante para la tarea que deba realizar el sistema. En MEXICA las estructuras de conocimiento contienen información para generar una historia. En la investigación actual las estructuras de conocimiento del sistema contienen información para componer una melodía. En el sistema producto de este trabajo de investigación, los ejemplos son melodías generadas por compositores humanos; podemos llamar a estos ejemplos el *repertorio* del sistema.

En la sección 4.3 se describen de forma detallada las estructuras de conocimiento y la forma de obtenerlas a partir de los ejemplos previos que el usuario provee al sistema.

3.2.2. El ciclo *E-R*

El modelo *engagement-reflection* consta de dos partes: una de generación y una de evaluación. Durante la parte generativa el sistema emplea las estructuras que se encuentran en la base de conocimiento para construir una cadena de ideas. La asociación de estas ideas se logra utilizando las constricciones con las que trabaja el modelo. Durante la parte de evaluación el

sistema revisa el material generado y lo modifica de ser necesario.

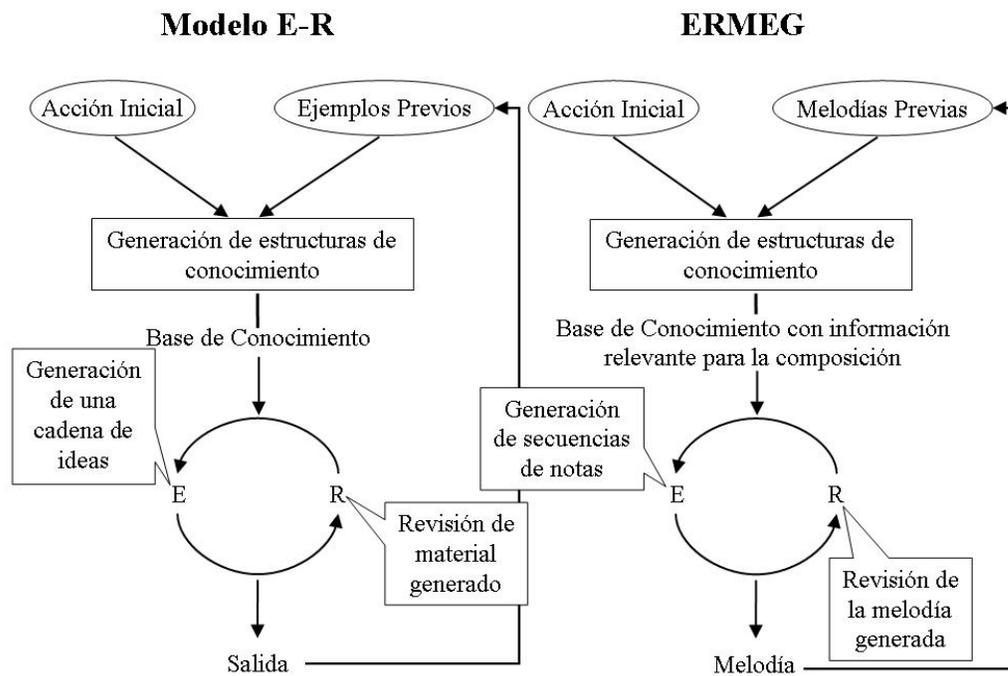
En el actual trabajo de investigación se utiliza el modelo *Engagement-Reflection* en un programa de computadora para la composición de melodías musicales. El sistema tiene el nombre de ERMEG (*Engagement-Reflection MEloodies Generator*). Esta idea surge porque la música y la composición musical no tienen significado semántico; pero se habla siempre de las emociones que producen. Aún así, los intentos de modelar el proceso de composición musical empleando la computadora, han utilizado en muy pocas ocasiones la componente emocional de la música. Además, por la característica de asociación de ideas basadas en vínculos emocionales del modelo E-R, este parece ser adecuado para la composición musical que carece de significado semántico, y en lugar de eso tiene significado emocional.

ERMEG genera melodías monofónicas. Durante *engagement* el sistema produce secuencias de notas musicales guiado por constricciones. Durante *reflection* el sistema: 1) rompe puntos muertos (*impasses*) generados durante *engagement*; 2) verifica la coherencia de la melodía en progreso y si es necesario, inserta notas musicales para satisfacerla; 3) evalúa la novedad del material producido. Como resultado de la evaluación y modificaciones hechas, *reflection* genera un conjunto de directrices que sirven como constricciones durante *engagement*. En la figura 3.2 se muestra un esquema donde se observa cómo se corresponden las principales características del modelo *engagement-reflection* con las partes del diseño del sistema compositor de melodías: ERMEG.

3.3. Tensión y Emoción

El sistema MEXICA utiliza vínculos emocionales y de tensiones entre los personajes. En este trabajo no es posible utilizar la misma representación de emociones, pues aquella corresponde a la interacción entre personajes. En ERMEG no hay personajes. Existen notas musicales, conjuntos de ellas que forman frases musicales y conjuntos de frases que forman melodías. Siguiendo las ideas de Gelernter, [Gelernter, 1994] “las emociones son el pegamento que une a las ideas”, y utilizando este pegamento con el modelo *engagement-reflection* para unir ideas musicales, desarrollamos una representación computacional de emociones y tensiones que pueda ser empleada por el sistema compositor de melodías.

En la literatura científica se pueden encontrar trabajos sobre música y emo-



, *Figura 3.2: Mapeo de las características del modelo E-R y las características del sistema ERMEG*

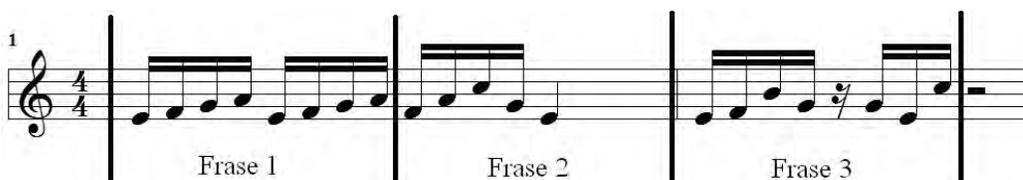
ciones. Los investigadores han tratado de averiguar cuáles son las características en la música que nos hacen percibirla de cierto modo. Decir que una pieza musical es alegre, es asignarle una etiqueta, pero ¿qué es lo que nos hace percibirla así? ¿cuáles son las características que tiene una pieza alegre que no tiene una que no lo es? Hay muchos parámetros para juzgar una pieza así. Son tantos, que no es fácil incluirlos todos en un sistema de cómputo. Aún si se hiciera reflejaría solo una verdad parcial, pues lo que para algunos es alegre para otros es triste.

En este trabajo se han tomado en cuenta algunas características musicales para construir un programa que pueda componer melodías basándose en una representación computacional de emociones. Las características extraídas de la música que el sistema utiliza han sido seleccionadas después de revisar la literatura acerca de música y emociones. Se menciona en trabajos científicos por ejemplo: que el tempo influye fuertemente en el tipo de emociones que se perciben en las piezas musicales; tempos rápidos tienden a evocar felicidad y tempos lentos tienden a evocar tristeza [Dalla Bella, Peretz, Rousseau & Gosselin, 2001]. Khalfa et al., [Khalifa, Roy, Rainville, Dalla Bella, & Peretz, 2008] han hecho experimentos en los que encontraron que ni el tempo ni el ritmo solos pueden generar emociones. Lo anterior se logra cuando se combina el ritmo y la melodía. Juslin y Sloboda, [Juslin & Sloboda, 2001] han mostrado que en la música tonal tonos altos y tempos rápidos generan felicidad; variaciones grandes de tono (saltos) y tonos altos pueden generar ansiedad; tonos graves y tempos lentos pueden generar tristeza. De los resultados mencionados se puede enunciar lo siguiente:

- La velocidad de la ejecución de una obra musical (ritmo), es una característica importante para distinguir emociones en la música.
- Las variaciones en una melodía musical (i.e. cambios entre las notas que la forman), influyen también en la percepción de emociones. El registro tonal en que se desarrolla la melodía; los cambios de una nota a otra; o los cambios a los que conduce un pasaje musical (i.e. como inicia y como termina), constituyen el tipo de variaciones a las que nos referiremos a lo largo de este trabajo. Hay otras variaciones cuyo impacto también es significativo (i.e. volumen, matices, etc.), que no consideramos para simplificar el diseño.
- La combinación del ritmo y las variaciones de la melodía son un fac-



, *Figura 3.3: Melodía de ejemplo.*



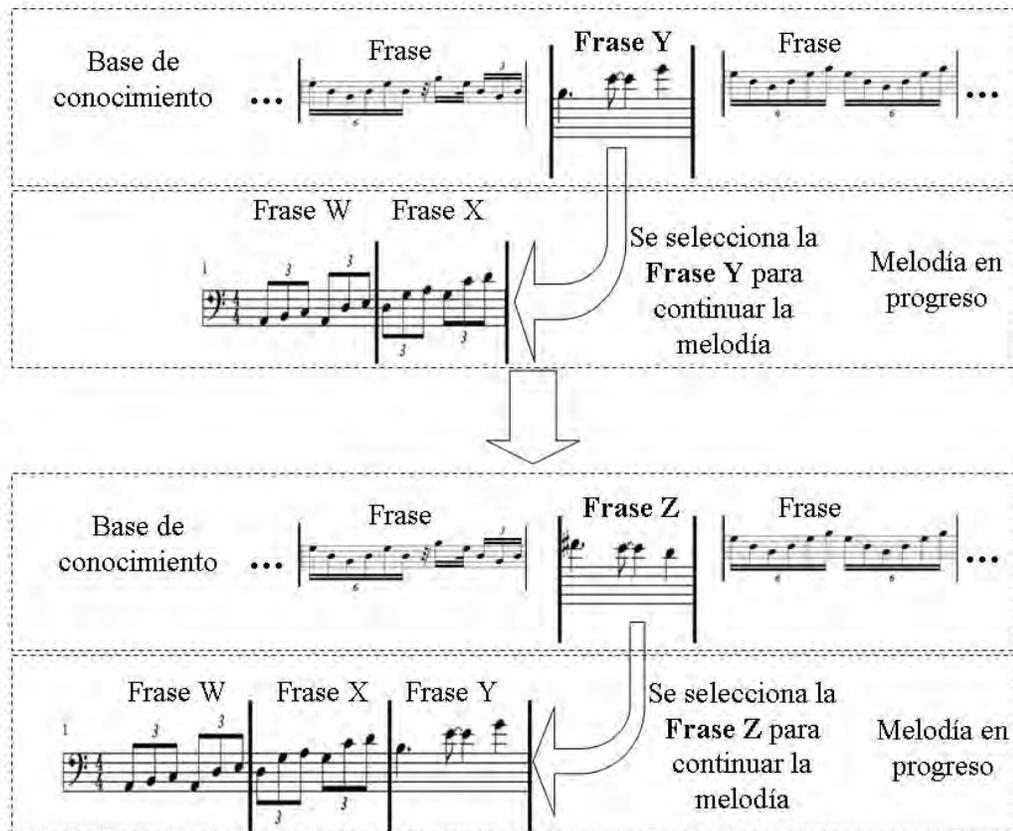
, *Figura 3.4: Melodía de ejemplo subdividida en frases composicionales. Las líneas verticales indican la división en la melodía.*

tor importante para determinar algún estado emocional producido por la melodía. Cualquiera de las dos características (ritmo y variaciones) aisladas no tienen el efecto que tienen combinadas.

Con base en todo lo anterior, en este trabajo de investigación calcularemos un valor de emoción y tensión para un grupo de notas. Esto es equivalente con los vínculos emocionales entre personajes que utiliza MEXICA. Para las representaciones computacionales de emociones y tensiones de grupos de notas, el programa emplea un tipo de emoción que es un continuo entre alegría y tristeza y un tipo de tensión que es un continuo entre ansiedad y calma. El sistema calcula los valores numéricos de tensión y emoción por medio de un sistema difuso.

Las melodías previas son arbitrariamente divididas en grupos de notas que se denominan frases composicionales (FC). Por lo tanto, una melodía se define como un conjunto de frases composicionales. En la figura 3.3 se muestra un ejemplo de melodía, y en la figura 3.4 la misma melodía dividida en frases. A cada una de estas frases se les calcula su valor de emoción y tensión. Esto es equivalente con los vínculos emocionales utilizados en MEXICA.

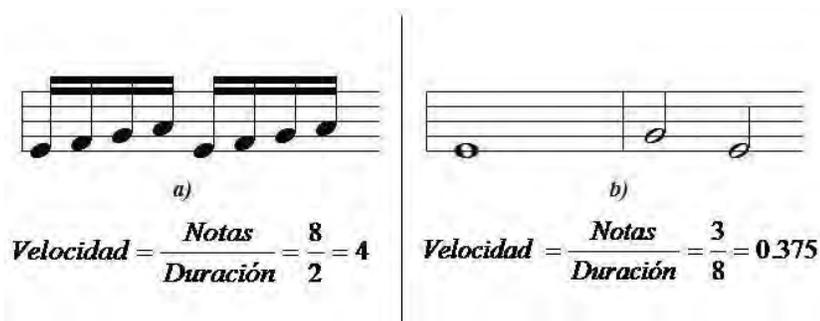
La base de datos del sistema contiene pequeñas frases composicionales que provienen de los ejemplos previos. El sistema genera secuencias de frases para obtener una melodía. En la figura 3.5 se muestra un diagrama que ejemplifica lo anterior.



, **Figura 3.5:** El sistema selecciona y agrega la Frase Y a la composición en progreso. Después, con la Frase Y incluida en la melodía en progreso, el sistema selecciona y agrega la Frase Z a la melodía.

Las características que nos sirven para calcular la representación de emoción y tensión en las melodías son:

- **Velocidad:** Cada frase tiene una velocidad asociada. Esta velocidad es la relación entre el número de notas y la duración de la frase (i.e. $\frac{\text{Notas}}{\text{Duración}}$). Cuando existe una relación Notas/Duración entre 0 y 2 se considera que la velocidad es baja; si la relación Notas/Duración está entre 0 y 4 se considera que la velocidad es mediana; cuando la relación Notas/Duración está entre 2 y 4 o mayor se considera que la velocidad es alta. Como estos rangos se traslapan utilizamos lógica difusa para resolver la incertidumbre, tanto en la velocidad como en las demás



, *Figura 3.6:* a) Una frase con muchas notas en una duración de 2 tiempos constituye una frase rápida, el valor obtenido ($Velocidad = 4$) corresponde a ese rango y b) Una frase con pocas notas en una duración de 8 tiempos constituye una frase lenta, el valor obtenido (0.375) corresponde a ese rango.

características. Estos valores son arbitrarios y surgen en base a la experiencia de este autor escuchando melodías lentas, de velocidad media y rápidas. A un mayor número de notas en una menor duración de frase se considera rápida. Si la frase es larga y tiene pocas notas es lenta. En la figura 3.6 se muestra un ejemplo de una frase rápida y una lenta. La frase a) obtiene un valor que cae dentro del rango de velocidad alta, En b) se obtiene un valor de velocidad de 0.375 que corresponde al rango de velocidad lenta¹.

- **Diferencia interválica:** La diferencia interválica está definida como la diferencia (en intervalos) que existe entre la primera y la última nota de una frase. Cuando existe un intervalo entre 0 y 6 tonos se considera que la diferencia es pequeña; cuando existe un intervalo que está entre 0 y 12 tonos se considera que la diferencia es mediana; cuando existe un intervalo entre 6 y 12 tonos o más se considera que la diferencia es grande². En la figura 3.7 se muestra un ejemplo con una amplia

¹El valor 0.375 cae dentro del rango de dos tipos de velocidad, baja y mediana. Para simplificar la explicación en este apartado solo se hace referencia de su pertenencia a uno de los rangos, pero en realidad pertenece a los dos. Lo mismo aplica para la demás características. Los detalles de cómo trabaja la lógica difusa para resolver estos traslapes se verán en la sección 4.2

²Los rangos se traslapan y se utiliza lógica difusa para resolver los traslapes. El detalle se muestra en la sección 4.2



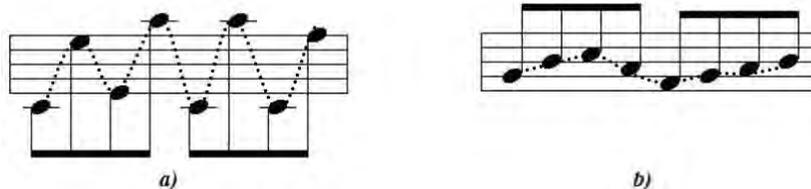
, *Figura 3.7:* a) Una frase que inicia con una nota y que termina con otra $10\frac{1}{2}$ intervalos por encima de la primera tiene una diferencia interválica grande. b) Una frase que inicia con una nota y que termina con otra $3\frac{1}{2}$ intervalos por encima de la primera tiene una diferencia interválica media.

diferencia interválica y un ejemplo con una diferencia interválica media.

- **Variación interválica:** Esta característica mide las variaciones entre las notas de una frase. Ella nos dice si la frase está cambiando abruptamente, o suave y gradualmente de nota a nota. Por ejemplo; cuánto cambia de la primer nota a la segunda; cuánto cambia de la segunda a la tercera y cuánto de la tercera a la cuarta. El promedio de los cambios entre todas las notas es la variación interválica. Si existe un promedio de cambio entre 0 y 1 tonos se considera que la variación es baja (pequeña); si está entre 0.5 y 3.5 tonos la variación es mediana; si cae entre 1 y 6 o más la variación es grande³. Si la variación promedio de cambio es pequeña los cambios entre notas son suaves; si es grande los cambios son abruptos. En la figura 3.8 se muestra un ejemplo.
- **Octava promedio:** Resultados de investigaciones muestran que registros musicales altos tienden a generar ansiedad y registros bajos tienden a generar calma. La octava promedio de la melodía nos indica el registro que tiene un cierto pasaje musical. Esta característica nos sirve para encontrar tensiones (ansiedad-calma) en las melodías.

Los niveles de tensión y emoción obtenidos por ERMEG son valores numéricos. Ellos provendrán de un cálculo hecho por un sistema difuso que se verá en la sección 4.2. El sistema difuso utiliza los valores numéricos de las características mencionadas. A continuación se desarrolla una explicación detallada de cómo se calculan los valores numéricos de cada una de estas características.

³Los rangos se traslapan y se utiliza lógica difusa para resolver los traslapes. El detalle se muestra en la sección 4.2



, **Figura 3.8:** En a) los cambios entre nota y nota son grandes, el promedio de variación es de 8.928 tonos, corresponde con el rango de variación grande; En b) los cambios entre las notas son pequeños, el promedio de variación es de 1.285 tonos, este corresponde con la variación mediana.

3.3.1. Velocidad

La Velocidad es una relación entre la cantidad de notas que tiene una FC y el tiempo en el que transcurren. Para ejemplificar, supóngase una FC de duración de 2 compases de 4/4 con tres notas en ella. La relación de notas y duración ($\frac{3 \text{ Notas}}{8 \text{ Tiempos}}$) determina que la frase es de velocidad baja (ver sección 3.3, página 46: Velocidad). Podemos pensar en frases cuyas notas tengan duraciones de 3, 3 y 2 tiempos o 4, 2 y 2 tiempos, tales frases tienen la misma densidad de notas en el mismo espacio de tiempo y por tanto ambas constituyen una frase lenta. La figura 3.9 muestra un ejemplo.

La velocidad baja no es un problema grave, pues podemos partir de velocidades de cero, es decir, frases que no tienen notas pero si tienen duración.

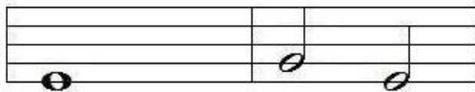
$$\frac{0 \text{ Notas}}{\text{Cualquier Duración} \neq 0}$$

Para la velocidad rápida debemos encontrar un límite superior, pues teóricamente podríamos añadir más y más notas a una frase haciéndola infinitamente más rápida, ya que su longitud no cambiaría.

$$\frac{\infty \text{ Notas}}{\text{Duración} \neq 0}$$

El autor ha establecido el límite superior en 4 notas por unidad de tiempo (*beat*). Para mostrar con un ejemplo, supóngase una frase de 2 compases de 4/4 con 32 notas en ella. La figura 3.10 muestra el ejemplo. Tal frase es considerada una frase rápida (ver página 46: Velocidad), y cualquier frase con una velocidad por encima de esta lo será también. Con los ejemplos de

las figuras 3.9 y 3.10, es fácil notar la distinción entre las frase lenta y rápida, pues tienen la misma duración y si se ejecutaran simultáneamente al final una de ellas tocó 3 notas y la otra 32; lo cual es una enorme diferencia.



, **Figura 3.9:** Ejemplo de una frase lenta con duraciones de notas de $\frac{1}{2}$, $\frac{1}{4}$ y $\frac{1}{4}$ tiempos. **Velocidad = 0.375**



, **Figura 3.10:** Ejemplo de una frase rápida con duraciones de notas de $\frac{1}{8}$ de tiempo. **Velocidad = 4.0**

Conocer los límites superior e inferior no es todo, pues no todas las frases musicales tendrán la misma forma que los ejemplos. Algunas tendrán números intermedios de notas y no podrían ser consideradas solo lentas o rápidas. Por tal motivo se ha utilizado la lógica difusa en este sistema para poder tratar con la incertidumbre que de otro modo se generaría.

3.3.2. Diferencia Interválica

Siguiendo los resultados de estudios que se han hecho en el campo de la música y emociones [Juslin & Sloboda, 2001], los fragmentos musicales que se tocan en registros altos favorecen respuestas emocionales como la alegría o ansiedad por las altas frecuencias que generan; mientras que fragmentos que se ejecutan en registros más bajos favorecen respuestas como tristeza o tranquilidad por las bajas frecuencias que generan.

La diferencia interválica es la distancia (en intervalos) que existe entre la primera nota y la última de una frase composicional. Ella nos dice que ha habido un cambio que está llevando a la frase musical a diferentes registros tonales. Por ejemplo: si una frase musical tiene una diferencia interválica grande (ver sección 3.3, página 47: Diferencia Interválica), quiere decir que se ha alejado mucho de la nota con la que inició. Si la frase inició en una



, **Figura 3.11:** Ejemplo de una frase con diferencia interválica grande ($10\frac{1}{2}$ tonos).

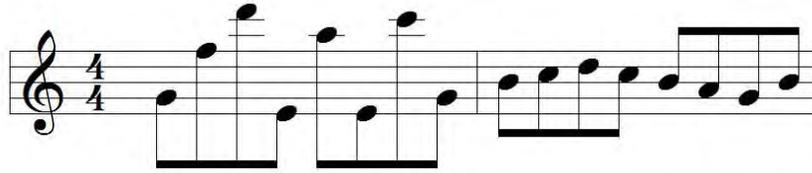
nota grave indica que la frase terminó en una nota aguda, o en una nota más grave que la inicial. La frase ha alcanzado diferentes registros tonales. Es por eso que esta característica es importante, pues funciona como indicador de la dirección emocional que puede tomar una melodía. En la figura 3.11 se muestra un ejemplo.

El autor ha establecido subjetivamente lo que es una diferencia grande o pequeña. La diferencia pequeña es definida de la misma forma que la velocidad cero, es decir, sin ninguna diferencia (la primer y última nota son la misma). La diferencia grande se ha establecido en 12 tonos de distancia (2 octavas), y cualquier frase que sea mayor que eso será considerada de diferencia grande (ver sección 3.3, página 47: Diferencia Interválica).

3.3.3. Variación Interválica

Esta característica contempla los cambios, medidos en intervalos, que existen entre las notas de una frase musical. Es importante porque puede suceder que una frase empiece y termine con la misma nota, con lo cual no habría ninguna diferencia interválica y por lo tanto podríamos equivocarnos al decir que la frase no cambia su representación emocional. La variación interválica mide el promedio de los cambios entre todas las notas de una FC.

Supongamos por ejemplo que hay dos FC consecutivas; ninguna de ellas tiene diferencia interválica, pero la primera de ellas tiene cambios grandes y la segunda tiene cambios pequeños (ver sección 3.3, página 48: Variación Interválica) entre nota y nota. En la figura 3.12 se muestra el ejemplo mencionado. Siguiendo los resultados de experimentos en estudios sobre música [Juslin & Sloboda, 2001], las variaciones melódicas amplias provocan ansiedad, tensión, alegría, etc., mientras que variaciones moderadas o cortas provocan calma, tranquilidad, tristeza, etc. Entonces, en el ejemplo citado habría una transición de un estado emocional ansioso o alegre a uno calmado o triste. El autor ha definido subjetivamente que la variación interválica grande entre notas en una melodía es 6 tonos (1 octava); cualquier variación mayor también

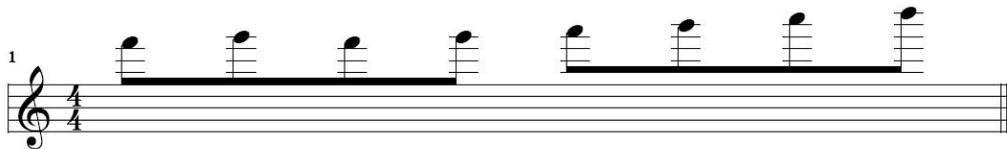


, *Figura 3.12: Ejemplo de dos frases con distinto tipo de variación interválica*

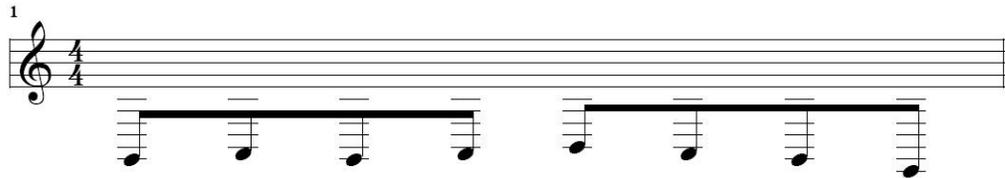
será considerada grande. La mínima variación como en el caso anterior es la no variación.

3.3.4. La octava promedio

Esta característica nos indica el registro promedio en el que se toca una frase composicional. Este registro es de interés, pues tiene importantes efectos en la percepción musical. Estudios sobre música y emociones [Juslin & Sloboda, 2001], muestran que los registros altos tienden a generar sensaciones de ansiedad y que melodías tocadas en registros bajos tienden a generar sensaciones de calma. Por lo tanto, mientras mayor sea la octava promedio, mayor tensión genera; mientras menor sea la octava promedio, menor tensión genera. En la figura 3.13 se muestra un ejemplo de octava promedio alta. En la figura 3.14 se muestra un ejemplo de octava promedio baja.



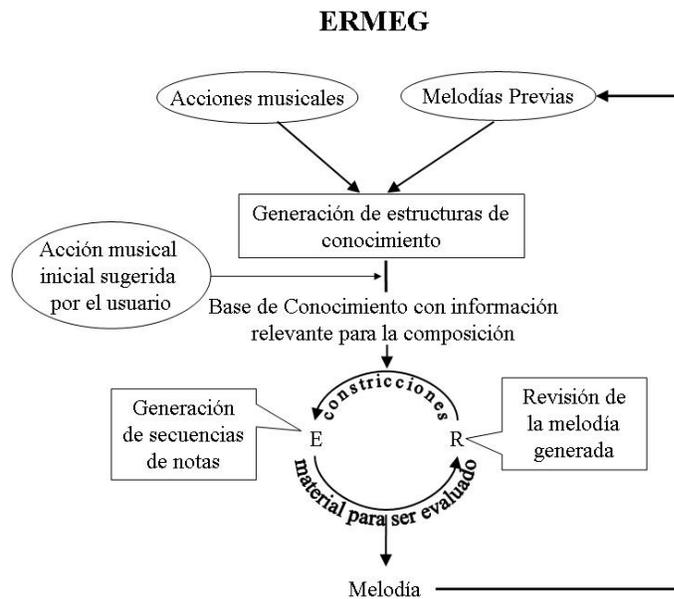
, *Figura 3.13: Ejemplo de frase de octava promedio alta, la octava promedio es 7.5*



, *Figura 3.14: Ejemplo de frase de octava promedio baja, la octava promedio es 3*

3.4. Discusión

El modelo computacional *engagement-reflection* ha sido utilizado antes en MEXICA y ha tenido resultados notables. Se ha empleado utilizando vínculos emocionales entre personajes para guiar la generación de una historia. No obstante, los vínculos utilizados en MEXICA no sirven en un sistema de cómputo para la generación de melodías. Por tal motivo, ha sido necesario construir una representación de emociones y tensiones para un sistema compositor.



, *Figura 3.15: Esquema General de ERMEG*

En este capítulo se han mostrado las características que se utilizan para construir una representación emocional. Esto, para ser utilizado en un sistema compositor de melodías. En la figura 3.15 se muestra un diagrama general de ERMEG. Las partes que lo forman serán analizadas con detalle en el siguiente capítulo. También se verá con detalle la forma en que se construye y utiliza la representación de emociones con el modelo E-R para guiar la composición de una melodía musical.

Capítulo 4

Descripción del Sistema

El objetivo de este capítulo es describir el sistema compositor de melodías producto de la investigación de este trabajo de tesis. El capítulo está dividido en cuatro partes que se enumeran a continuación:

- La primer parte corresponde a la introducción del sistema compositor de melodías basado en el modelo E-R producto de este trabajo de investigación.
- Para componer una melodía el sistema calcula niveles numéricos de emociones y tensiones en las frases composicionales mediante el uso de lógica difusa. En la segunda parte del capítulo se aborda el diseño del sistema difuso para tal efecto. Con los niveles numéricos de emociones y tensiones el sistema construye estructuras de conocimiento necesarias para la composición.
- La tercera parte describe la construcción de la base de conocimiento que emplea el sistema durante el proceso de composición.
- La cuarta parte corresponde a la descripción del ciclo *Engagement-Reflection*. Al final de la ejecución del ciclo E-R el sistema produce una melodía musical.

4.1. Introducción

El campo de la generación de música por medio de una computadora ha dado lugar al desarrollo de varios trabajos, algunos de ellos con resultados notables. Las arquitecturas de este tipo de sistemas son tan diversas como los métodos que emplean para generar melodías: redes neuronales, algoritmos genéticos, métodos simbólicos, etc. Pero todas ellas comparten una característica: utilizan la teoría musical como elemento principal en el proceso de composición y normalmente no consideran otros factores relevantes que intervienen en la elaboración de una melodía.

Personas sin profundos conocimientos musicales son capaces de componer música. La teoría ayuda mucho a esta labor, más no es un impedimento para aquellos que no la tienen. ¿Pero, cómo componer sin teoría musical? Los recursos empleados en estos casos son: la experiencia en la ejecución de algún tipo de música, de algún instrumento, las emociones que transmiten las distintas partes de una melodía y la forma de interconectar dichas partes. En este trabajo de investigación se ha desarrollado un programa de computadora para la composición musical llamado ERMEG. Este sistema no está basado en reglas musicales estrictas, sino en representaciones del contenido emocional de la música, el cual es un aspecto que no se ha tomado en cuenta en muchos de los trabajos realizados en composición automática.

ERMEG utiliza lógica borrosa o difusa para obtener una representación computacional del contenido emocional. Esta representación proporciona información necesaria para que el sistema asigne una etiqueta emocional a las frases composicionales.

Estructuras de conocimiento necesarias para el proceso de composición son generadas. Éstas, pueden verse como la experiencia de un compositor humano, que le ayudan a saber cómo continuar una composición.

ERMEG genera una melodía nueva como resultado de ciclos de interacción entre los estados *engagement* y *reflection*. En dichos ciclos utiliza la información obtenida del sistema difuso y las estructuras de conocimiento que posee.

El sistema es controlado por una cantidad de parámetros cuyos valores son determinados por el usuario y que controlan el comportamiento del sistema. El detalle de funcionamiento de ERMEG se muestra a continuación.

4.2. Sistema Difuso

Las variables de entrada al sistema difuso son: Velocidad, Diferencia Interválica y Variación Interválica. Las salidas son los valores numéricos de Tensión y Emoción. En la figura 4.1 se muestra el esquema del diseño conceptual del sistema difuso.

El valor numérico de tensión se multiplica por la octava promedio, pues mientras mayor sea la octava promedio, mayor tensión genera y mientras menor sea la octava promedio, menor tensión genera.

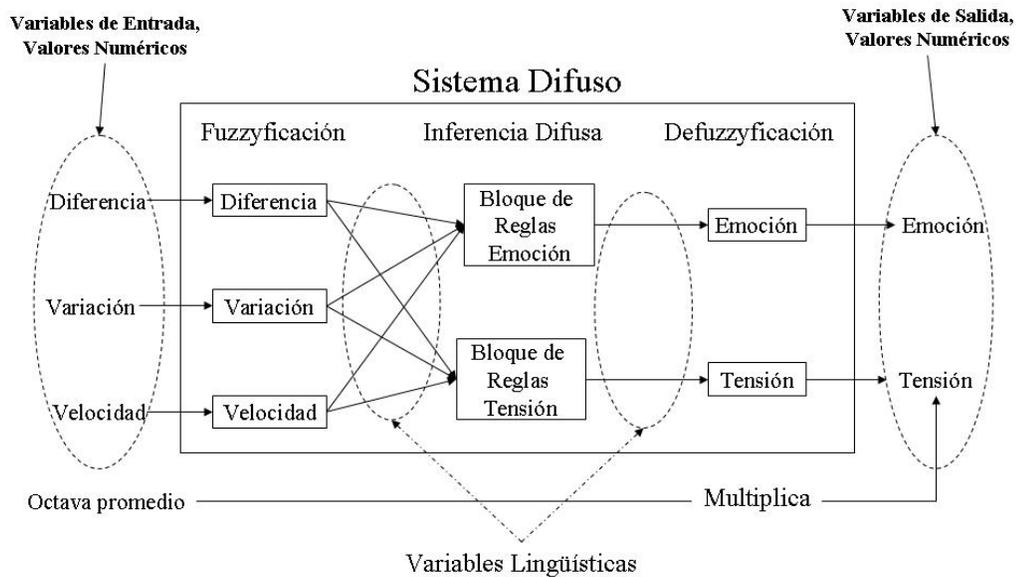
Los pasos que sigue el sistema difuso son los siguientes:

1. Se convierten los valores numéricos de velocidad, diferencia y variación interválica en valores difusos (fuzzificación). *Una variable numérica de entrada se convierte en N variables lingüísticas con diferentes valores de membresía.*
2. Se lleva a cabo la inferencia difusa que tiene lugar en los bloques de reglas. Esta, consiste en la aplicación de reglas difusas utilizando las variables lingüísticas de la fuzzificación.
3. La salida de los bloques de reglas son variables lingüísticas.
4. La desfuzzificación en la salida traduce las variables lingüísticas en variables análogas o numéricas.

En la tabla 4.1 se muestran las estadísticas para el sistema difuso empleado en esta tesis.

VARIABLES DE ENTRADA	3
VARIABLES DE SALIDA	2
BLOQUES DE REGLAS	2
REGLAS	54
FUNCIONES DE MEMBRESÍA	19

, **Tabla 4.1:** Estadísticas del Proyecto

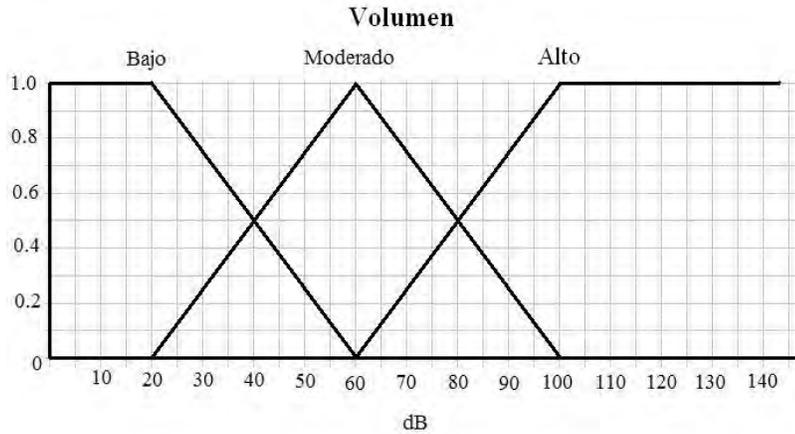


, *Figura 4.1: Diseño Conceptual del Sistema Difuso*

4.2.1. Variables Lingüísticas

En esta sección se detalla la definición de las variables lingüísticas y de las funciones de membresía. Llamamos variable lingüística a aquella noción o concepto que vamos a calificar de forma difusa. Por ejemplo: la altura, el volumen, la edad, la velocidad. . . Le aplicamos el adjetivo “lingüística” porque definiremos sus características mediante el lenguaje hablado. Las variables lingüísticas se usan para traducir valores reales (analógicos) en lingüísticos. Los posibles valores de una variable lingüística no son números sino que son llamados términos lingüísticos. La función de membresía es aquella aplicación que asocia a cada elemento de un conjunto difuso el grado con que pertenece al valor lingüístico asociado. Los conjuntos difusos son caracterizados por sus funciones de membresía.

Por ejemplo: El lector puede imaginar una orquesta interpretando una melodía. El volumen de la orquesta puede ser bajo, moderado o alto. Cada uno de ellos lo describe mejor o peor dependiendo del actual nivel de volumen. Cada término es definido por una función de membresía (FM) que especifica para cualquier valor de la variable de entrada el correspondiente grado de membresía. Las funciones de membresía de los términos de una variable lin-



, *Figura 4.2: Funciones de Membresía de “Volumen”*

Volumen Bajo	con grado de membresía de 0.9
Volumen Moderado	con grado de membresía de 0.1
Volumen Alto	con grado de membresía de 0.0

, *Tabla 4.2: Grados de membresía a los términos bajo, moderado y alto para un nivel de volumen de 25 dB.*

güística son usualmente representadas en una gráfica. La figura 4.2 muestra la gráfica del ejemplo.

En la tabla 4.2 se muestran los grados de membresía de los términos bajo, moderado y alto para un nivel de volumen de 25 dB.

Las variables lingüísticas tienen que ser definidas para todas las variables de entrada y salida del sistema.

Entradas

Para el diseño de las variables lingüísticas, empezamos por definir el universo de discurso, conjunto universal o referencial. Llamamos universo de discurso al rango de valores que pueden tomar los elementos que poseen la propiedad expresada por la variable lingüística. Para el caso de la velocidad de una FC este universo va desde 0 hasta 4. El límite 4, de una frase rápida (ver sección 3.3, página 46: Velocidad), se calcula de una frase de 32 notas con una duración de 2 compases (ver fig. 3.10). La relación de $\frac{\text{Notas}}{\text{Duración}}$ tiene un

#	Nombre de la Variable	Valor Mín. del Universo de Discurso	Valor Máx. del Universo de Discurso	Términos
1	Diferencia	0	12	Baja Media Alta
2	Variación	0	6	Baja Media Alta
3	Velocidad	0	4	Baja Media Alta

, **Tabla 4.3:** Entradas del Sistema difuso

resultado de 4.

Para la diferencia interválica se mencionó antes (ver sección 3.3, página 47: Diferencia) que la máxima diferencia que se acepta es de 12 tonos. Ése es el límite superior del universo de discurso y cualquier diferencia mayor obtendrá la misma evaluación lingüística. El límite inferior, igual que en el caso anterior, es 0. Este valor representa que no hay diferencia.

La variación interválica tiene un universo de discurso que tiene como límite inferior cero y como límite superior el promedio de las máximas variaciones posibles en una frase (ver sección 3.3, página 48: Variación). Si todas las notas están separadas por la máxima diferencia que separa dos notas consecutivas (6 tonos) el promedio de sus diferencias es 6 tonos. Este es el límite superior de la diferencia interválica y representa una variación interválica alta.

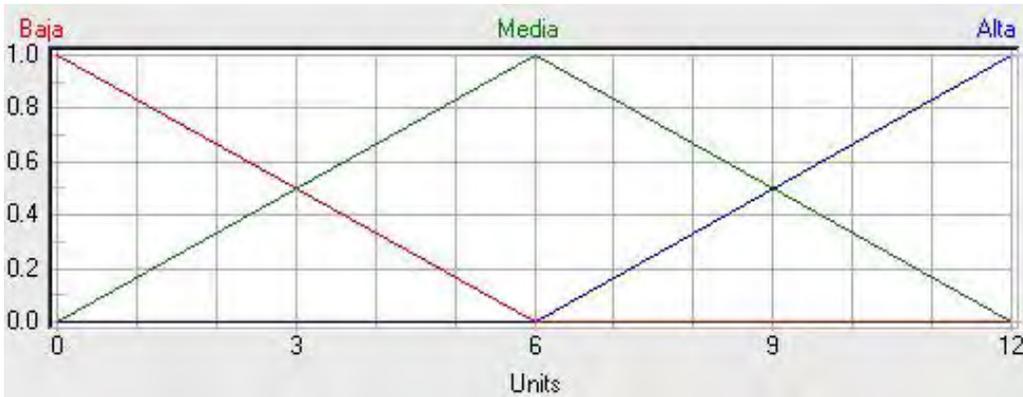
En la tabla 4.3 se puede observar lo antes mencionado. También se muestran los términos lingüísticos que se incluyen en cada uno de los universos de discurso de las variables.

A continuación se detallan los diagramas de cada variable de entrada.

En la figura 4.3, y en la tabla 4.4 se muestran las características de la variable Diferencia Interválica; las funciones de membresía, sus formas, puntos de definición y el universo de discurso.

En la figura 4.4, y en la tabla 4.5 se muestran las características de la variable Variación Interválica; las funciones de membresía, sus formas, puntos de definición y el universo de discurso.

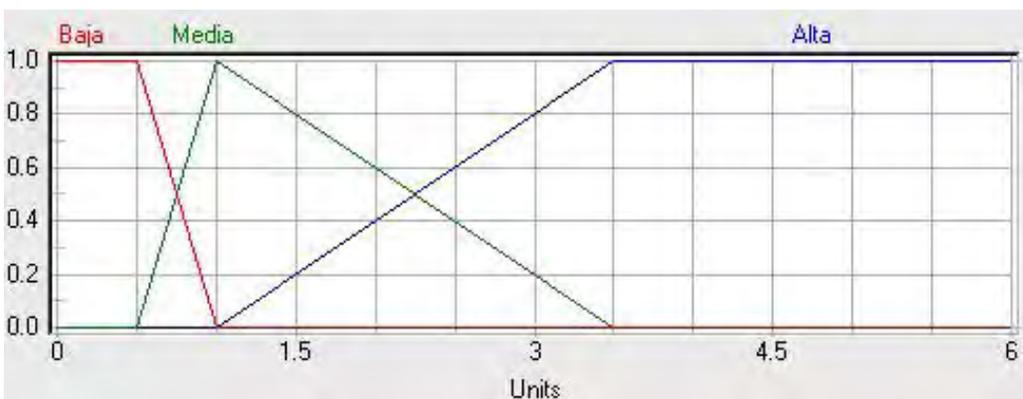
En la figura 4.5, y en la tabla 4.6 se muestran las características de la variable Velocidad; las funciones de membresía, sus formas, puntos de definición y el



, **Figura 4.3:** Funciones de Membresía de “Diferencia”

Nombre del término	Forma de los segmentos	Puntos de Definición (x, y)		
Baja	Lineal	(0, 1)	(6, 0)	(12, 0)
Media	Lineal	(0, 0)	(6, 1)	(12, 0)
Alta	Lineal	(0, 0)	(6, 0)	(12, 1)

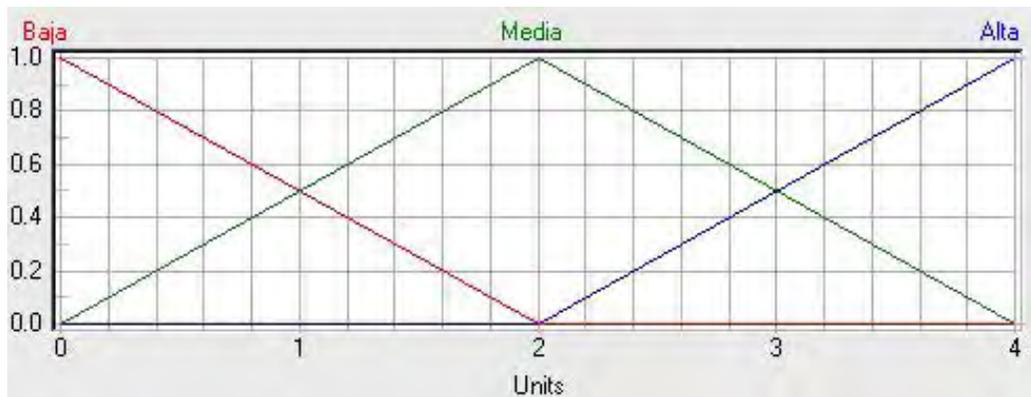
, **Tabla 4.4:** Puntos de definición de las funciones de membresía de “Diferencia”



, **Figura 4.4:** Funciones de Membresía de “Variación”

Nombre del término	Forma de los segmentos	Puntos de Definición (x, y)		
Baja	Lineal	(0, 1)	(0.5, 1)	(1, 0)
Media	Lineal	(0, 0)	(0.5, 0)	(1, 1)
Alta	Lineal	(0, 0)	(1, 0)	(3.5, 1)

, **Tabla 4.5:** Puntos de definición de las funciones de membresía de “Variación”



, **Figura 4.5:** Funciones de Membresía de “Velocidad”

Nombre del término	Forma de los segmentos	Puntos de Definición (x, y)		
Baja	Lineal	(0, 1)	(2, 0)	(4, 0)
Media	Lineal	(0, 0)	(2, 1)	(4, 0)
Alta	Lineal	(0, 0)	(2, 0)	(4, 1)

, **Tabla 4.6:** Puntos de definición de las funciones de membresía de “Velocidad”

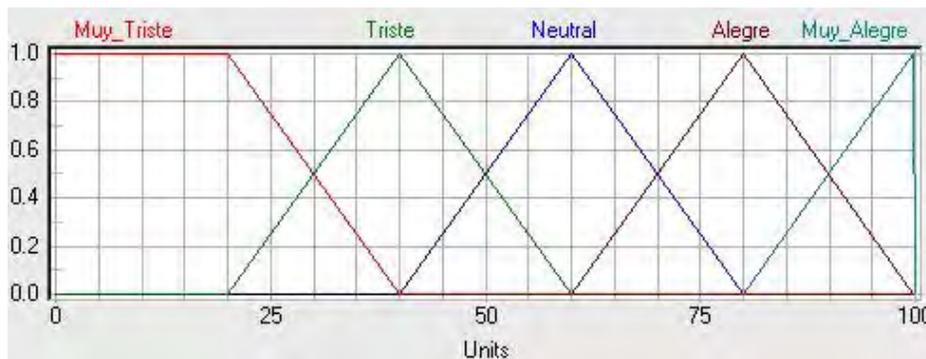
universo de discurso.

Salidas

Para las variables de salida, los universos de discursos corren en un rango total de 0 % a 100 % en la variables de tensión de emoción. En la tabla 4.7 se muestra la figura y también los términos lingüísticos que aparecen para cada

#	Nombre de la Variable	Valor Mín. del Universo de Discurso	Valor Máx. del Universo de Discurso	Términos
4	Emoción	0	100	Muy_Triste Triste Neutral Alegre Muy_Alegre
5	Tensión	0	100	Muy_Calmado Calmado Neutral Ansioso Muy_Ansioso

, *Tabla 4.7: Salidas del Sistema difuso*



, *Figura 4.6: Funciones de Membresía de "Emoción"*

una de ellas.

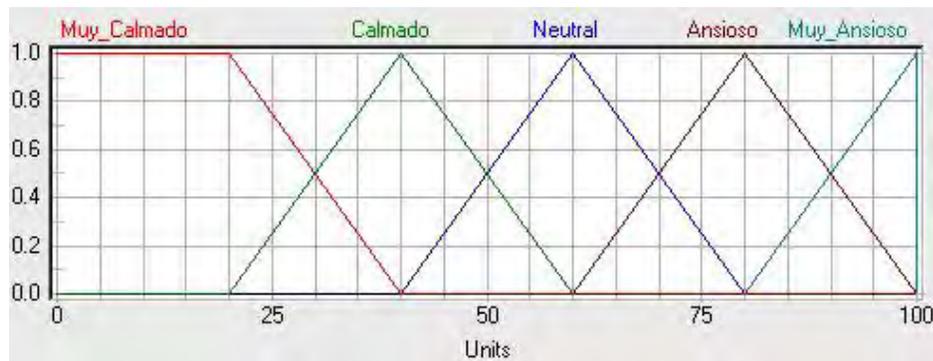
A continuación se detallan los diagramas de cada variable de salida.

En la figura 4.6, y en la tabla 4.8 se muestran las características de la variable Emoción; las funciones de membresía, sus formas, puntos de definición y el universo de discurso.

En la figura 4.7, y en la tabla 4.9 se muestran las características de la variable Tensión; las funciones de membresía, sus formas, puntos de definición y el universo de discurso.

Nombre del término	Forma de los segmentos	Puntos de Definición (x, y)		
Muy_ Triste	Lineal	(0, 1)	(20, 1)	(40, 0)
		(100, 0)		
Triste	Lineal	(0, 0)	(20, 0)	(40, 1)
		(60, 0)	(100, 0)	
Neutral	Lineal	(0, 0)	(40, 0)	(60, 1)
		(80, 1)	(100, 1)	
Alegre	Lineal	(0, 0)	(60, 0)	(80, 1)
		(100, 0)		
Muy_ Alegre	Lineal	(0, 0)	(80, 0)	(100, 1)

, **Tabla 4.8:** Puntos de definición de las funciones de membresía de “Emoción”



, **Figura 4.7:** Funciones de Membresía de “Tensión”

Nombre del término	Forma de los segmentos	Puntos de Definición (x, y)		
Muy_ Calmado	Lineal	(0, 1)	(20, 1)	(40, 0)
		(100, 0)		
Calmado	Lineal	(0, 0)	(20, 0)	(40, 1)
		(60, 0)	(100, 0)	
Neutral	Lineal	(0, 0)	(40, 0)	(60, 1)
		(80, 1)	(100, 1)	
Ansioso	Lineal	(0, 0)	(60, 0)	(80, 1)
		(100, 0)		
Muy_ Ansioso	Lineal	(0, 0)	(80, 0)	(100, 1)

, **Tabla 4.9:** Puntos de definición de las funciones de membresía de “Tensión”

4.2.2. Bloques de Reglas

Los bloques de reglas contienen la estrategia de control de un sistema de lógica difusa. La parte del “*IF*” describe la situación para la cual las reglas han sido diseñadas. La parte del “*THEN*” describe la respuesta del sistema difuso para una determinada situación.

Para construir los bloques de reglas se han utilizado algunos resultados de investigaciones en el campo de emociones y música [Juslin & Sloboda, 2001] [Dalla Bella, Peretz, Rousseau & Gosselin, 2001] [Khalfa, Roy, Rainville, Dalla Bella, & Peretz, 2008]. Con ellos ha sido posible definir las relaciones entre las variables de entrada Velocidad, Diferencia Interválica, Variación Interválica, Octava Promedio y las variables de salida Tensión y Emoción. Como resultado se ha obtenido el siguiente conjunto de reglas:

Emoción:

1. IF Velocidad = Baja AND Diferencia = Baja AND Variación = Baja THEN Emoción = Muy_Triste
2. IF Velocidad = Baja AND Diferencia = Baja AND Variación = Media THEN Emoción = Muy_Triste
3. IF Velocidad = Baja AND Diferencia = Baja AND Variación = Alta THEN Emoción = Muy_Triste
4. IF Velocidad = Baja AND Diferencia = Media AND Variación = Baja THEN Emoción = Muy_Triste
5. IF Velocidad = Baja AND Diferencia = Media AND Variación = Media THEN Emoción = Muy_Triste
6. IF Velocidad = Baja AND Diferencia = Media AND Variación = Alta THEN Emoción = Muy_Triste
7. IF Velocidad = Baja AND Diferencia = Alta AND Variación = Baja THEN Emoción = Triste
8. IF Velocidad = Baja AND Diferencia = Alta AND Variación = Media THEN Emoción = Triste
9. IF Velocidad = Baja AND Diferencia = Alta AND Variación = Alta THEN Emoción = Triste
10. IF Velocidad = Media AND Diferencia = Baja AND Variación = Baja THEN Emoción = Triste
11. IF Velocidad = Media AND Diferencia = Baja AND Variación = Media THEN Emoción = Neutral
12. IF Velocidad = Media AND Diferencia = Baja AND Variación = Alta THEN Emoción = Neutral
13. IF Velocidad = Media AND Diferencia = Media AND Variación = Baja THEN Emoción = Triste
14. IF Velocidad = Media AND Diferencia = Media AND Variación = Media THEN Emoción = Neutral
15. IF Velocidad = Media AND Diferencia = Media AND Variación = Alta THEN Emoción = Neutral
16. IF Velocidad = Media AND Diferencia = Alta AND Variación = Baja THEN Emoción = Neutral
17. IF Velocidad = Media AND Diferencia = Alta AND Variación = Media THEN Emoción = Neutral
18. IF Velocidad = Media AND Diferencia = Alta AND Variación = Alta THEN Emoción = Alegre
19. IF Velocidad = Alta AND Diferencia = Baja AND Variación = Baja THEN Emoción = Alegre

20. IF Velocidad = Alta AND Diferencia = Baja AND Variación = Media THEN Emoción = Alegre
21. IF Velocidad = Alta AND Diferencia = Baja AND Variación = Alta THEN Emoción = Muy_Alegre
22. IF Velocidad = Alta AND Diferencia = Media AND Variación = Baja THEN Emoción = Alegre
23. IF Velocidad = Alta AND Diferencia = Media AND Variación = Media THEN Emoción = Muy_Alegre
24. IF Velocidad = Alta AND Diferencia = Media AND Variación = Alta THEN Emoción = Muy_Alegre
25. IF Velocidad = Alta AND Diferencia = Alta AND Variación = Baja THEN Emoción = Muy_Alegre
26. IF Velocidad = Alta AND Diferencia = Alta AND Variación = Media THEN Emoción = Muy_Alegre
27. IF Velocidad = Alta AND Diferencia = Alta AND Variación = Alta THEN Emoción = Muy_Alegre

Tensión:

1. IF Velocidad = Baja AND Diferencia = Baja AND Variación = Baja THEN Emoción = Muy_Calmado
2. IF Velocidad = Baja AND Diferencia = Baja AND Variación = Media THEN Emoción = Muy_Calmado
3. IF Velocidad = Baja AND Diferencia = Baja AND Variación = Alta THEN Emoción = Muy_Calmado
4. IF Velocidad = Baja AND Diferencia = Media AND Variación = Baja THEN Emoción = Muy_Calmado
5. IF Velocidad = Baja AND Diferencia = Media AND Variación = Media THEN Emoción = Calmado
6. IF Velocidad = Baja AND Diferencia = Media AND Variación = Alta THEN Emoción = Calmado
7. IF Velocidad = Baja AND Diferencia = Alta AND Variación = Baja THEN Emoción = Calmado
8. IF Velocidad = Baja AND Diferencia = Alta AND Variación = Media THEN Emoción = Calmado
9. IF Velocidad = Baja AND Diferencia = Alta AND Variación = Alta THEN Emoción = Neutral
10. IF Velocidad = Media AND Diferencia = Baja AND Variación = Baja THEN Emoción = Muy_Calmado
11. IF Velocidad = Media AND Diferencia = Baja AND Variación = Media THEN Emoción = Muy_Calmado
12. IF Velocidad = Media AND Diferencia = Baja AND Variación = Alta THEN Emoción = Calmado
13. IF Velocidad = Media AND Diferencia = Media AND Variación = Baja THEN Emoción = Calmado
14. IF Velocidad = Media AND Diferencia = Media AND Variación = Media THEN Emoción = Neutral
15. IF Velocidad = Media AND Diferencia = Media AND Variación = Alta THEN Emoción = Neutral
16. IF Velocidad = Media AND Diferencia = Alta AND Variación = Baja THEN Emoción = Ansioso
17. IF Velocidad = Media AND Diferencia = Alta AND Variación = Media THEN Emoción = Ansioso
18. IF Velocidad = Media AND Diferencia = Alta AND Variación = Alta THEN Emoción = Muy_Ansioso
19. IF Velocidad = Alta AND Diferencia = Baja AND Variación = Baja THEN Emoción = Muy_Ansioso
20. IF Velocidad = Alta AND Diferencia = Baja AND Variación = Media THEN Emoción = Calmado
21. IF Velocidad = Alta AND Diferencia = Baja AND Variación = Alta THEN Emoción = Neutral
22. IF Velocidad = Alta AND Diferencia = Media AND Variación = Baja THEN Emoción = Neutral

23. IF Velocidad = Alta AND Diferencia = Media AND Variación = Media THEN Emoción = Neutral
24. IF Velocidad = Alta AND Diferencia = Media AND Variación = Alta THEN Emoción = Ansioso
25. IF Velocidad = Alta AND Diferencia = Alta AND Variación = Baja THEN Emoción = Ansioso
26. IF Velocidad = Alta AND Diferencia = Alta AND Variación = Media THEN Emoción = Muy_Ansioso
27. IF Velocidad = Alta AND Diferencia = Alta AND Variación = Alta THEN Emoción = Muy_Ansioso

Implicación

El operador de implicación utilizado para las reglas anteriores en ambos casos fue Mamdani. Se eligió este operador ya que se requería que el comportamiento fuera proporcional al cambio de la entrada numérica. Además se requería que fuera lento al aproximarse a los puntos diáfanos. De esta forma podemos mantenernos por más tiempo dentro de la zona de incertidumbre y con esto darle más opciones al sistema al momento de componer.

Agregación

El operador de agregación se eligió como consecuencia de elegir el operador de implicación Mamdani, pues su relación con el operador de agregación está dada desde la formación del operador de implicación. El operador de agregación elegido fue de unión.

Desfuzzyficación

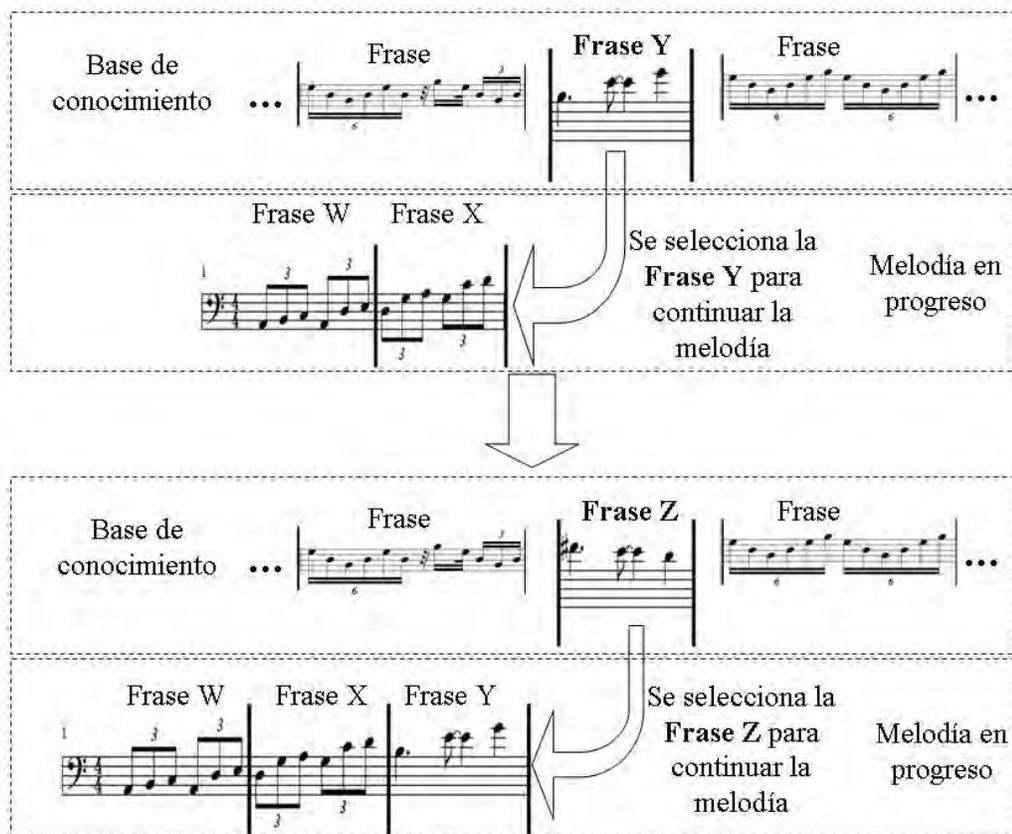
El método de desfuzzyficación empleado es el centro de área o gravedad (COA). Se usó porque es ampliamente utilizado. Además porque el cálculo puede hacerse rápidamente.

4.3. Construcción de la Base de Conocimiento

El sistema crea las estructuras de conocimiento necesarias para el proceso de composición. Éstas pueden verse como la experiencia de una persona, que le ayudan a saber cómo continuar una composición. Esta experiencia es extraída de un conjunto llamado “Melodías Previas” que el usuario provee al sistema. Ellas conforman lo que podríamos llamar el repertorio del sistema.

4.3.1. División de las Melodías

La unidad básica de composición del sistema son las Frases Composicionales (ver página 45). Es decir, ERMEG compone melodías generando secuencias de FC. (ver figura 4.8). Estas frases las obtiene de un conjunto de melodías que el usuario provee. En esta sección se explica cómo se obtienen las FC.



, *Figura 4.8: ERMEG compone melodías generando secuencias de Frases Composicionales. El sistema agrega la Frase Y y posteriormente la Frase Z a la melodía en progreso.*

Es necesario hacer un análisis de las melodías previas para poder extraer información relevante para el proceso de composición. El primer paso para este análisis es la división de las melodías previas en frases composicionales. La razón para emplear frases, las cuales son pequeños fragmentos musicales,

es que usualmente los compositores utilizan pequeños fragmentos que conocen, que han escuchado, ejecutado, y que sirven para ensamblar una melodía. Una característica importante acerca de las frases musicales que usualmente se emplean es que tienen una duración regular. Por regular nos referimos a duraciones que caen dentro de los estándares de la música. Con ellas podemos percibir en una melodía un sentido rítmico estable. Duraciones regulares son frases de un compás, de tres tiempos, de dos tiempos de duración, etc.

Para nosotros es importante encontrar los acentos de una melodía. Éstos confieren lo que podemos llamar sentido rítmico. Considerando la teoría musical y lo que naturalmente podemos percibir en la música, estos acentos se pueden encontrar en partes definidas. Por ejemplo: en una melodía con compases de $4/4$, podemos encontrar un acento en el primer tiempo del compás. Este momento es muy importante, pues es el acento más fuerte y distinguible. También podemos encontrar un acento en el tercer tiempo del compás. Este también es importante, pero es más débil. La posición de los acentos depende de la subdivisión de los compases. Para este trabajo de investigación las melodías que el sistema puede procesar y componer tienen subdivisiones de $4/4$. El sistema, por tanto, divide las melodías previas en frases composicionales de longitudes convenientes, haciéndolo entre los acentos de los compases, es decir, que subdivide en frases de dos tiempos, cuatro tiempos, seis tiempos, etc.

Las frases, a pesar de que se indique al sistema el tamaño de subdivisión conveniente, pueden ser de diferente longitud. Esto puede pasar porque una nota puede extenderse más allá de esa longitud y no sería normal cortar la nota. Para solucionar el problema la frase se extiende. En la figura 4.9 se muestra un ejemplo de cómo se subdividen las melodías previas en frases composicionales. Las líneas verticales indican la división de una frase. En la figura las FC no tienen la misma longitud pues a pesar de tener una duración de frase conveniente (2 tiempos) hay notas que se extienden más allá de esa duración. Las notas no se cortan, y por lo tanto la frase se extiende.

El tamaño de las frases puede tener efectos en las composiciones generadas. Por ejemplo, en composiciones del sistema que utilizan frases cortas, es difícil distinguir cuáles fueron usadas. Es decir, no sabremos de que melodía previa provienen. Si son largas en algunos casos podremos saber con certeza de que melodía previa fueron tomadas. Además de esto hay otro efecto que se puede notar con FC de distintos tamaños: si son largas, las melodías compuestas tienden a ser estables, musicalmente hablando, pues estamos tomando secciones grandes de melodías que estaban bien construidas. Si son pequeñas se



, **Figura 4.9:** Divisiones de frases. Las frases uno y dos tienen longitudes de dos tiempos. En las frases tres, cuatro y cinco las duraciones son de cuatro tiempos debido a la presencia de notas que se extienden más allá de los dos tiempos.

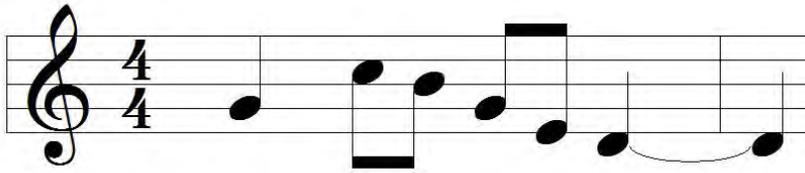
pierde la estabilidad, pues la frase contiene muy pocas notas, esto implica que se pierda significado musical de la frase, en el extremo, frases de una sola nota carecen de él. En general se busca que las frases sean de un tamaño apropiado. En experimentos hechos con el sistema se han obtenido buenos resultados con frases de dos tiempos.

Una vez que las melodías previas han sido divididas en frases podemos extraer información relevante para el proceso de composición. Cada frase tiene las características de velocidad, variación interválica y diferencia interválica que se mencionaron en la sección 3.3 del capítulo 3. Con estas características calculamos valores de tensión y emoción que generan.

También se construyen gráficas de la melodía y del ritmo de las melodías previas. Estas representaciones corresponden a las notas que tiene una melodía y el ritmo con que son ejecutadas. Estas estructuras sirven para comprobar que el proceso de composición está construyendo una melodía nueva con respecto a las que tiene en la base de conocimiento, o si el sistema está copiando las melodías previas en una composición. Comparando las gráficas es cómo el sistema concluye si las melodías son copias o no. Esto es importante, pues la intención del sistema es generar melodías nuevas con respecto a su base de conocimiento.

Para construir la gráfica de la melodía el sistema utiliza los valores MIDI de las notas. Para el ritmo las duraciones de las notas tendrán valores positivos y los silencios valores negativos. En la figura 4.10 se muestra un ejemplo para calcular las gráficas de melodía y ritmo que se muestran en las figuras 4.11 y 4.12 respectivamente. En la tabla 4.10 se muestra el listado de notas y duraciones¹.

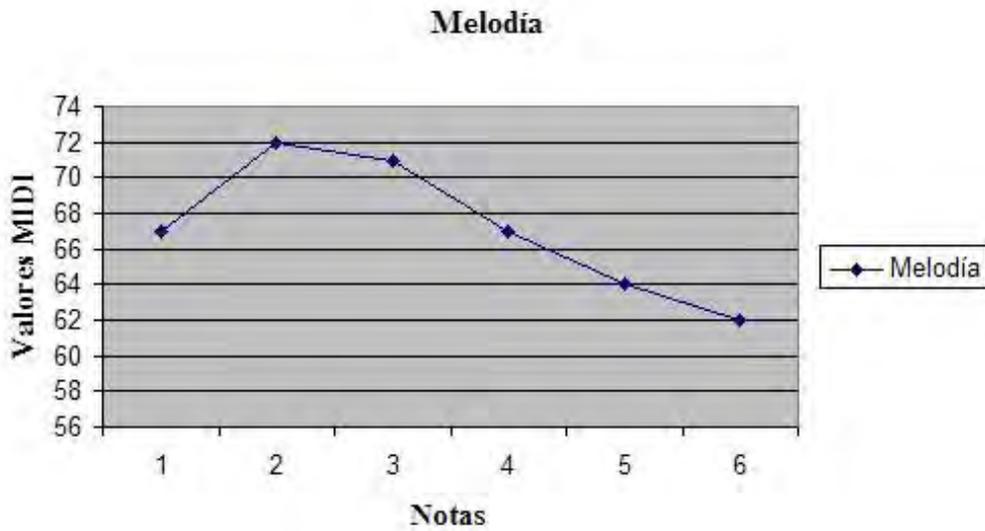
¹Las duraciones también son valores del formato MIDI. 1:000 corresponde a 960tics. Tic es la unidad de medida de tiempo en el formato MIDI.



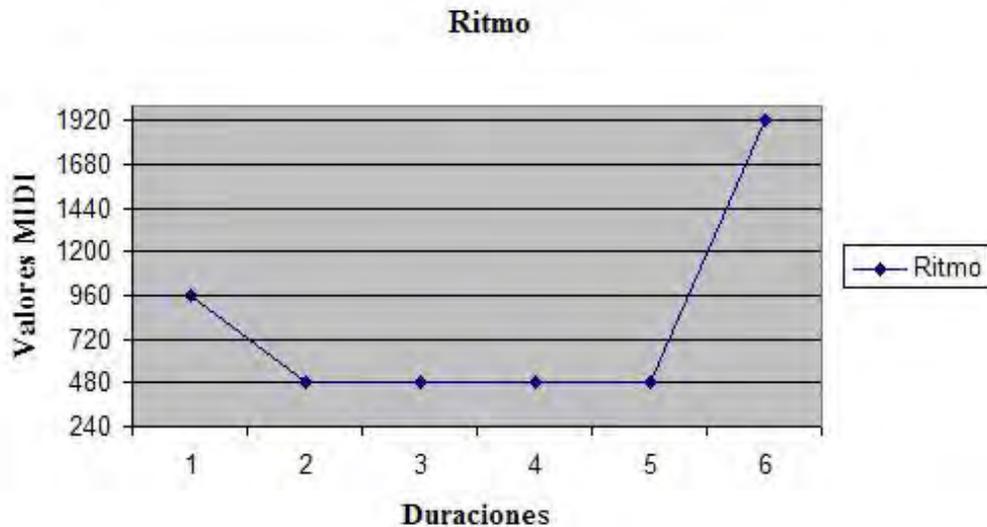
, *Figura 4.10: Breve melodía*

Nota MIDI	Duración MIDI
67	1:000
72	480
71	480
67	480
64	480
62	2:000

, *Tabla 4.10: Listado de notas y duraciones de la melodía del ejemplo*



, *Figura 4.11: Gráfica de Melodía del ejemplo.*



, *Figura 4.12: Gráfica de Ritmo del ejemplo*

Una vez que se han obtenido los valores de emoción y tensión de las frases los utilizamos para construir estructuras de conocimiento necesarias para llevar a cabo el proceso de composición.

4.3.2. Contextos

En ERMEG una composición previa se define como una sucesión de frases composicionales.

$$\text{Melodía_Previa} = \text{FC1} + \text{FC2} + \text{FC3} + \dots + \text{FCn}$$

Cada una de las frases composicionales (FC) de la melodía previa tiene asociados valores de emoción y tensión a los que nos referimos como valores locales de tensión y emoción. Pero desde la FC2 en adelante, a cada FC se le pueden asignar valores de tensión y emoción históricos. Los valores de tensión y emoción históricos son un vector que contiene todos los valores de tensión y emoción locales de las FC que preceden a otra FC. De esta forma por ejemplo, los valores históricos de emoción para la FC3 son [emoción local de FC1, emoción local de FC2]. En la figura 4.13 se ilustra el ejemplo.

<p>Contexto de FC3</p> <p>Locales</p> <p>Emoción: Emoción-FC3</p> <p>Tensión: Tensión-FC3</p> <p>Históricos</p> <p>Emoción: [Emoción-FC1, Emoción-FC2]</p> <p>Tensión: [Tensión-FC1, Tensión-FC2]</p>
--

, *Figura 4.13: Contexto FC3*

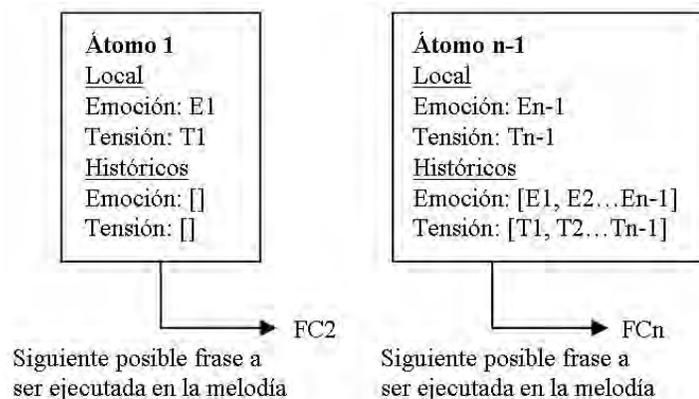
Melodía Previa = FC1 +	FC2 + ...	FCn			
<table border="1"> <tr> <td> <p>Contexto 1</p> <p><u>Local</u></p> <p>Emoción: E1</p> <p>Tensión: T1</p> <p><u>Histórico</u></p> <p>Emoción: []</p> <p>Tensión: []</p> </td> </tr> </table>	<p>Contexto 1</p> <p><u>Local</u></p> <p>Emoción: E1</p> <p>Tensión: T1</p> <p><u>Histórico</u></p> <p>Emoción: []</p> <p>Tensión: []</p>	<table border="1"> <tr> <td> <p>Contexto 2</p> <p><u>Local</u></p> <p>Emoción: E2</p> <p>Tensión: T2</p> <p><u>Histórico</u></p> <p>Emoción: [E1]</p> <p>Tensión: [T1]</p> </td> </tr> </table>	<p>Contexto 2</p> <p><u>Local</u></p> <p>Emoción: E2</p> <p>Tensión: T2</p> <p><u>Histórico</u></p> <p>Emoción: [E1]</p> <p>Tensión: [T1]</p>	<table border="1"> <tr> <td> <p>Contexto n</p> <p><u>Local</u></p> <p>Emoción: En</p> <p>Tensión: Tn</p> <p><u>Histórico</u></p> <p>Emoción: [E1,E2,..,En-1]</p> <p>Tensión: [T1,T2,..,Tn-1]</p> </td> </tr> </table>	<p>Contexto n</p> <p><u>Local</u></p> <p>Emoción: En</p> <p>Tensión: Tn</p> <p><u>Histórico</u></p> <p>Emoción: [E1,E2,..,En-1]</p> <p>Tensión: [T1,T2,..,Tn-1]</p>
<p>Contexto 1</p> <p><u>Local</u></p> <p>Emoción: E1</p> <p>Tensión: T1</p> <p><u>Histórico</u></p> <p>Emoción: []</p> <p>Tensión: []</p>					
<p>Contexto 2</p> <p><u>Local</u></p> <p>Emoción: E2</p> <p>Tensión: T2</p> <p><u>Histórico</u></p> <p>Emoción: [E1]</p> <p>Tensión: [T1]</p>					
<p>Contexto n</p> <p><u>Local</u></p> <p>Emoción: En</p> <p>Tensión: Tn</p> <p><u>Histórico</u></p> <p>Emoción: [E1,E2,..,En-1]</p> <p>Tensión: [T1,T2,..,Tn-1]</p>					

, *Figura 4.14: Representación de una melodía previa y sus Contextos.*

Toda esta información se graba en una estructura llamada Contexto; de tal forma que el Contexto está formado de cuatro elementos: la emoción local, la tensión local, el vector de emociones históricas y el vector de tensiones históricas. Tenemos entonces que a la FC1 corresponde el Contexto1 y a la FC2 el Contexto2, etc. (ver figura 4.14)

El contexto nos muestra las condiciones en las que una frase sucede. Nos indica los niveles de tensión y emoción locales e históricos. Con esto sabemos, que para llegar al estado de emoción y tensión local de la frase a la que pertenece el contexto, una forma de hacerlo es siguiendo la historia de tensiones y emociones que tiene la frase.

Conocer el contexto en el que ocurre una frase nos indica cómo se ha llegado hasta ahí y de qué forma. Pero para componer necesitamos saber también qué debemos hacer a continuación. Lo que queremos es conocer lo que se hizo en tal situación (contexto) en una melodía previa, es decir, qué frase musical se empleó. Esa información después nos será útil para la composición.



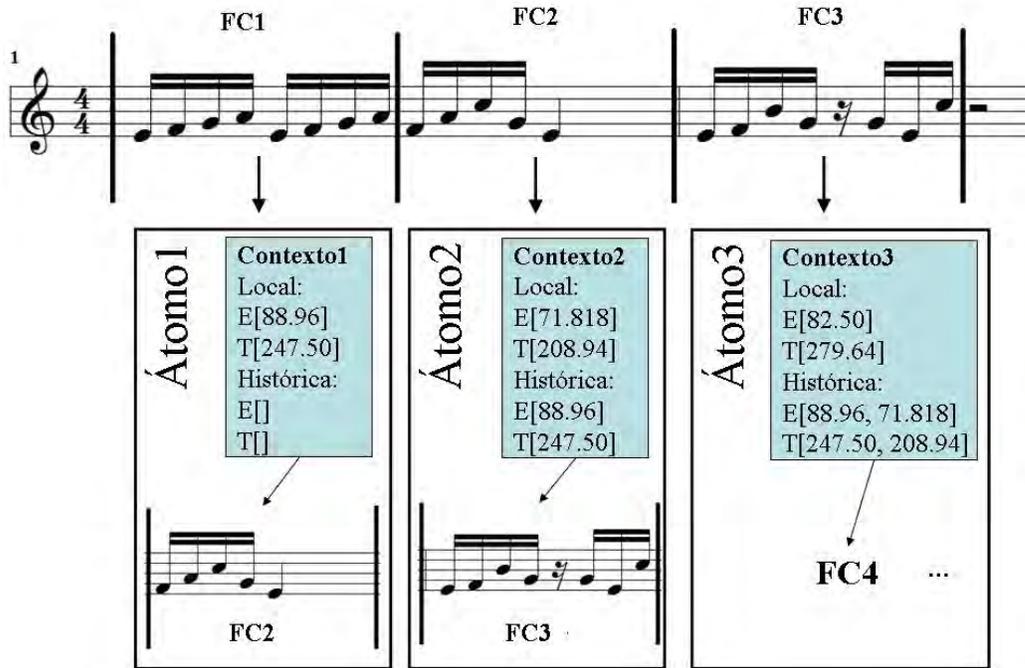
, *Figura 4.15: Dos átomos y sus FCs asociadas (o posibles frases a ejecutar).*

4.3.3. Átomos

El sistema vincula los Contextos con la siguiente frase composicional de la melodía. Así el Contexto1 es asociado a la frase composicional 2 (FC2), el Contexto2 es asociado a la FC3, y así sucesivamente (ver figura 4.15). El propósito de este tipo de asociaciones es establecer posibles caminos para continuar una composición dado un Contexto específico. De esta manera el sistema graba información del tipo “cuándo una melodía tiene un contexto como el Contexto1, una manera de continuar la composición es con la FC2”. Con esta información el sistema crea en memoria sus estructuras de conocimiento llamadas *átomos* que emplea durante el ciclo *engagement-reflection* para generar una nueva melodía. La idea central es asociar contextos (que representan el estado actual de la melodía en términos de nuestra representación de tensiones y emociones) con FCs a ser ejecutadas. En la figura 4.16 se muestra un ejemplo de la generación de contextos y átomos a partir de una secuencia de frases (melodía).

Para crear las estructuras de conocimiento el sistema lleva a cabo los siguientes pasos para cada una de las melodías previas. Utilizaremos la información de la figura 4.16:

1. El sistema inicializa el contexto y hace la primer frase en la melodía la actual frase composicional (*FC1*).
2. El sistema calcula sus valores de emoción y de tensión y actualiza el



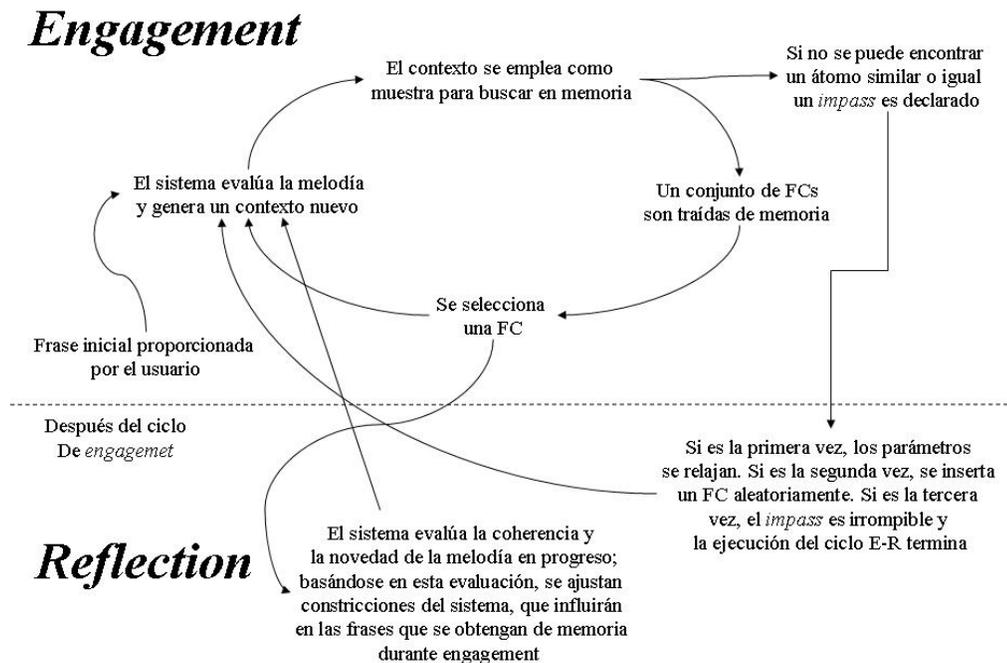
, **Figura 4.16:** Generación de contextos y átomos a partir de una secuencia de frases (melodía)

contexto (*Contexto1*).

3. El contenido del contexto es copiado a una nueva estructura llamada Átomo (*Átomo1*).
4. El sistema asocia la siguiente frase composicional (*FC2*) con el Átomo (*Átomo1*), nos referimos a esta frase (*FC2*) como la siguiente posible frase a ser ejecutada en la melodía.
5. El sistema toma la siguiente FC en la melodía (*FC2*) y la hace la actual FC, regresa al paso 2 y repite el proceso hasta que todas las frases en la melodía hayan sido procesadas.

4.4. Ciclo *Engagement-Reflection*

Durante la fase de generación, el sistema genera melodías como resultado de un ciclo *engagement-reflection* (ver figura 4.17). El número de repeticiones del ciclo completo E-R (parámetro *Ciclo ER*²) es un parámetro definido por el usuario.



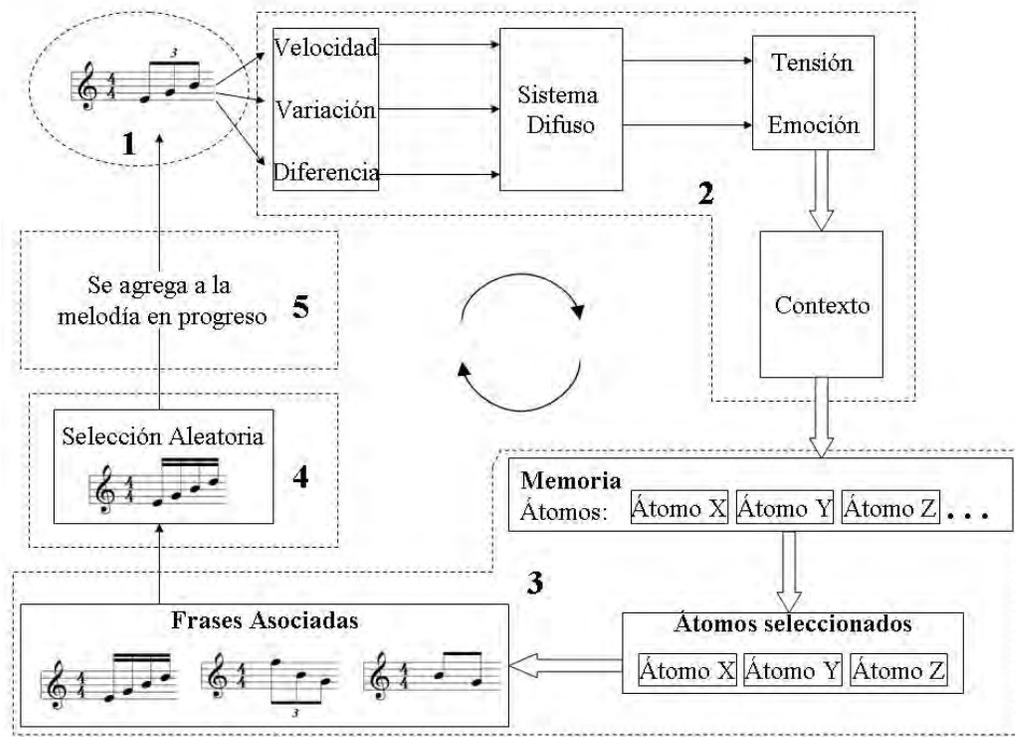
, Figura 4.17: Ciclo *Engagement-Reflection*

4.4.1. *Engagement*

Engagement es un ciclo cuyo número de repeticiones es un parámetro del sistema definido por el usuario (parámetro *Ciclo E*). El ciclo *engagement* sigue los siguientes pasos:

1. El usuario da al sistema una frase musical inicial.
2. El sistema calcula los valores de tensión y emoción de dicha frase y actualiza el Contexto.

²En la sección 4.5 se puede encontrar un resumen de los parámetros del sistema.



, *Figura 4.18: Pasos del ciclo engagement*

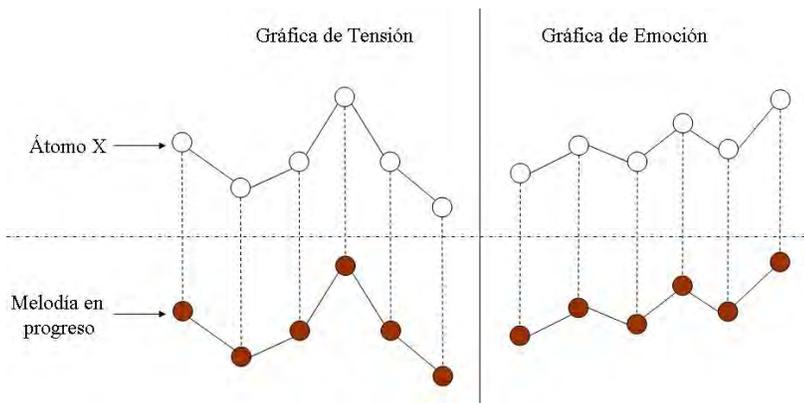
3. El contexto es empleado como muestra para localizar en memoria átomos que sean iguales o similares al contexto y obtener las FC asociadas.
4. El sistema selecciona aleatoriamente una de las FC obtenidas como la siguiente frase a ser ejecutada en la melodía en progreso.
5. La FC es agregada a la melodía en progreso, El Contexto se actualiza y el ciclo inicia otra vez (ir al paso 3).

En la figura 4.18 se muestra un ejemplo del ciclo *Engagement*. Los números en cada región indican el paso que se ejecuta.

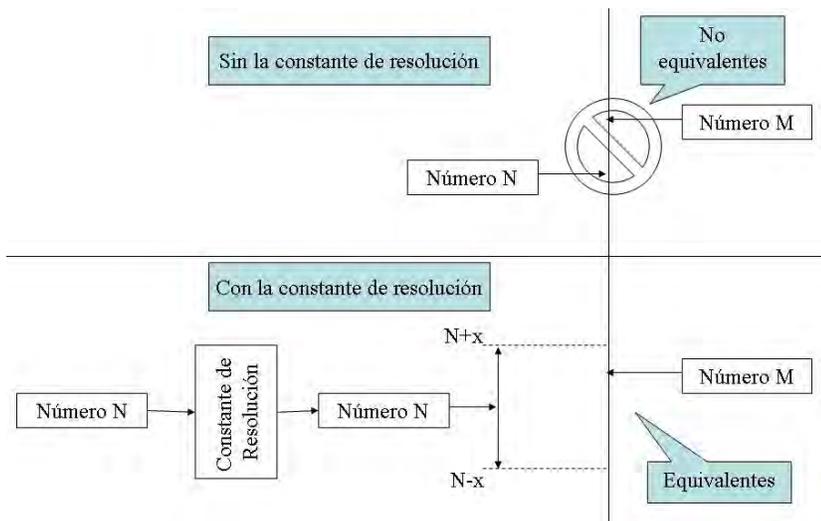
Como se muestra en la figura 4.18 el sistema actualiza el contexto de la melodía en desarrollo. Este contexto se utiliza para buscar en la memoria del sistema átomos similares o iguales. Un conjunto de frases composicionales son traídas de la memoria, y una de ellas es seleccionada para continuar la melodía en progreso.

Puede suceder también que el sistema no encuentre ningún contexto similar o igual. Cuando esto pasa se declara un punto muerto (*impass*), y es necesario ejecutar algunas rutinas para tratar de romper este bloqueo. Estas rutinas son llevadas a cabo por *reflection*.

Para agregar nuevas frases a la melodía en progreso es necesario distinguir cuando un contexto es similar o igual al contexto actual. En la figura 4.19 se muestra un *match* perfecto (iguales) entre un átomo y la melodía en progreso. Durante *engagement* el sistema emplea dos importantes parámetros para decidir si un contexto es igual o similar al átomo: la constante de resolución y la constante ACAS.



, *Figura 4.19: Match Perfecto entre un átomo y la melodía en progreso.*



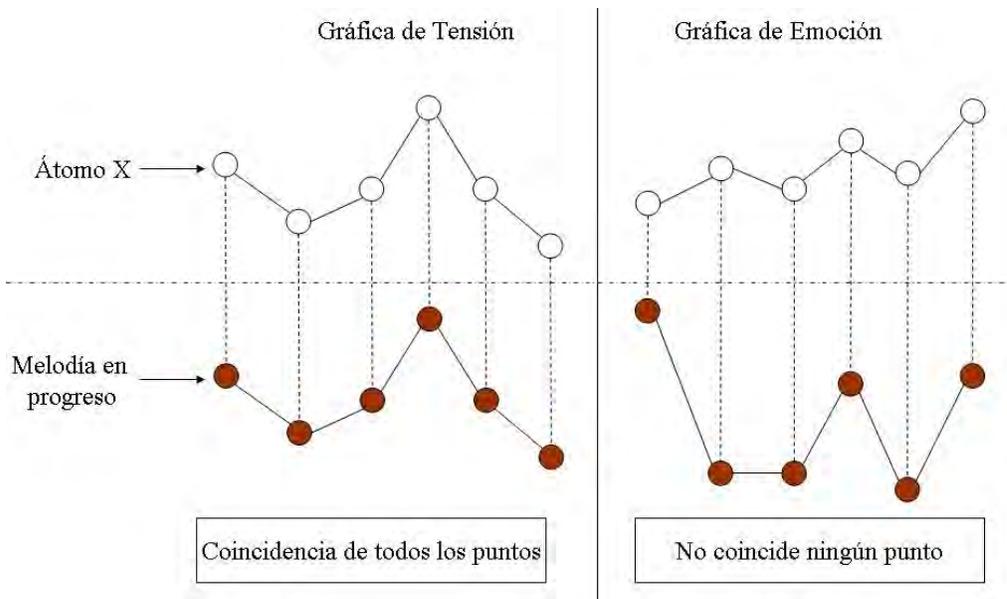
, *Figura 4.20: Equivalencia de valores empleando la constante de resolución*

La constante de resolución es un rango empleado para determinar si dos valores numéricos son considerados como equivalentes. Por ejemplo, si a la constante de resolución se le asigna el valor de 90 %, cualquier valor 90 % similar o más es considerado como equivalente. En la figura 4.20 se muestra un ejemplo donde se compara si dos valores numéricos son equivalentes. La comparación se hace en dos partes, primero sin la constante de resolución donde cada número es diferente y no son equivalentes. Luego con la constante de resolución en donde el número N se convierte en un rango que va desde $N-x$ hasta $N+x$. En tal rango cae el número M y consideramos que el número N y M son equivalentes.

La constante de resolución consta de dos partes: la constante de resolución local y la constante de resolución global. La primera de ellas se enfoca en encontrar similitudes a nivel local del contexto, es decir, sólo en la tensión y emoción local sin tomar en cuenta el vector histórico de tensiones y emociones. La segunda se enfoca en encontrar similitudes globales en el contexto, es decir, que toma en cuenta el vector histórico de tensiones y emociones. Cada una de ellas tiene un valor independiente, que puede ser libremente modificado por el usuario. Si por ejemplo, el usuario pide que la resolución local sea del 100 %, y que la resolución global sea del 0 %, lo que sucede aquí es que el usuario está quitándole toda la importancia a la información global, y dándole toda la importancia a la información local. Con esto lo que sucede es que el sistema busca en memoria cualquier átomo que se 100 % igual en el contexto local, sin importar el vector histórico de tensiones y emociones. Si sucediera lo contrario el sistema buscaría en memoria todos los átomos que fueran 100 % iguales en el contexto global sin importar el contexto local.

La constante ACAS (*Atom Containing the Associative Structure*) indica el porcentaje mínimo en que deben ser equivalentes el Contexto de la melodía en progreso y los átomos para producir una asociación. Por ejemplo, si la constante ACAS tiene un valor del 50 %, al menos el 50 % del contexto debe ser equivalente al del átomo. En la figura 4.21 se muestra un ejemplo.

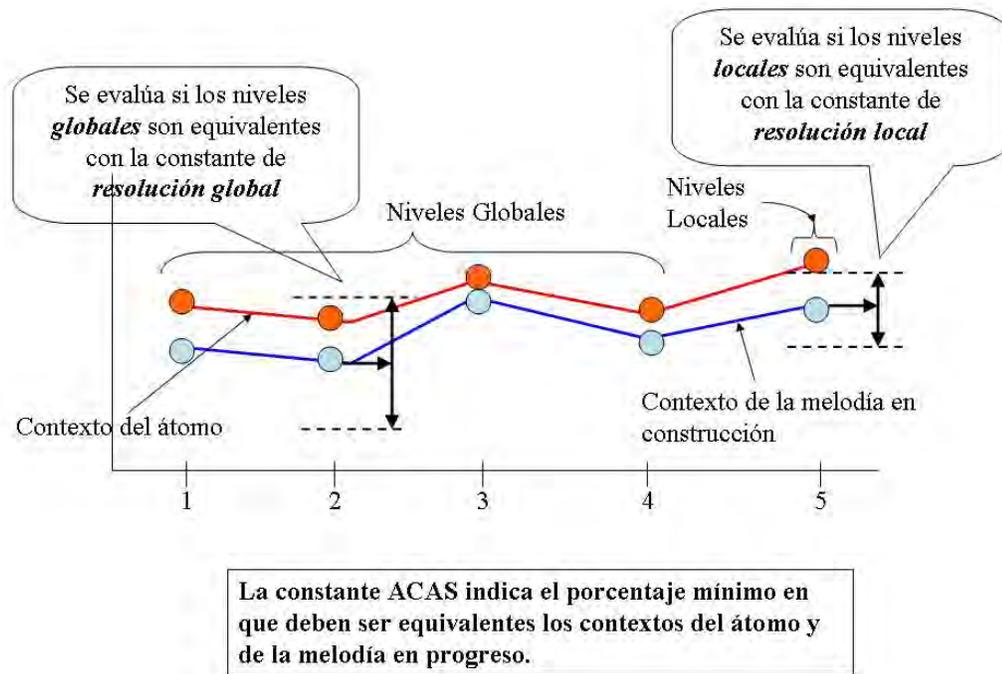
La constante ACAS consta de dos partes. Estas partes trabajan como complementarias. Una de ellas es la constante ACAS superior, y la otra es la constante ACAS inferior. Los valores por omisión de la constante ACAS superior e inferior son 100 % y 50 %. Estos parámetros al igual que las constantes de resolución pueden ser modificados libremente por el usuario. Con los valores por omisión el sistema busca encontrar contextos que sean entre 50 % (ACAS inferior) y 100 % (ACAS superior) similares. Si se modifica la constante ACAS inferior, y se ajusta su valor al 0 %, pedimos al sistema que



, **Figura 4.21:** El contexto es equivalente al átomo en 50 %, pues solo la mitad de los puntos comparados son equivalentes

seleccione todos los átomos del sistema (ACAS entre 0 % y 100 %), pues no nos importa si se parecen o no. Si modificamos la constante ACAS superior e inferior y las ajustamos al 50 % y 20 % respectivamente, pedimos al sistema que seleccione los átomos con similitud entre el 20 % y el 50 %.

En la figura 4.22 se muestra una imagen de las gráficas del contexto de un átomo y del contexto de la melodía en progreso. En la figura los puntos representan los valores de tensión o emoción de la melodía en progreso y del átomo. Ellos se comparan con un rango de similitud establecido por la constante de resolución: local y global. El número de coincidencias o equivalencias entre los puntos de la gráfica es calculado. Para poder decidir si el átomo es similar o igual, el porcentaje de coincidencias debe caer en el rango de la constante ACAS.



, **Figura 4.22:** Asociación de los contextos del átomo y de la melodía en progreso usando las constantes de resolución y ACAS

Las constantes de resolución y ACAS son parámetros que el usuario puede modificar libremente. La modificación de estos parámetros supone un cambio en la manera de encontrar contextos similares o iguales dentro del conjunto de los átomos en la memoria del sistema. Estos parámetros son muy importantes ya que ERMEG es un instrumento de investigación. Con ellos, los investigadores tienen mucha libertad para experimentar con el sistema. Los resultados del sistema modificando las constantes de resolución y ACAS pueden ser muy interesantes. Si las constantes toman los siguientes valores:

Resolución local	=	0 %
Resolución global	=	0 %
ACAS inferior	=	0 %
ACAS superior	=	100 %

El sistema dispondrá de todos los átomos que tenga en memoria y las melodías generadas podrían parecerse un poco “alocadas”.

Si las constantes toman los siguientes valores:

Resolución local	=	100 %
Resolución global	=	100 %
ACAS inferior	=	100 %
ACAS superior	=	100 %

El sistema tal vez no dispondrá de muchos átomos para extraer frases y continuar la composición de la melodía. En algunos casos tal vez no encuentre ningún átomo. También, dado que la selección de átomos es muy rigurosa podríamos encontrar que las melodías no son “alocadas”, sino todo lo contrario. Incluso podría suceder que por lo estricto de los parámetros el sistema no encuentre formas alternas de seguir una composición, y copie las melodías previas.

Con el valor de estos parámetros ajustados en niveles intermedios, se pueden encontrar resultados muy interesantes, que no suenan demasiado “alocados”, y que no son copias de las melodías previas que tiene el sistema.

Engagement finaliza cuando se llega a un punto muerto o cuando un número predefinido de ciclos ha sido ejecutado. El número de repeticiones del ciclo *engagement* es un parámetro definido por el usuario. Se declara un punto muerto cuando ya no se pueden agregar más frases a la melodía en progreso, pues no es posible encontrar ningún átomo similar o igual en la memoria con el contexto de la melodía.

4.4.2. *Reflection*

Durante *Reflection* el sistema trata de romper puntos muertos (*impasses*) y evalúa la coherencia y la novedad de la melodía en progreso.

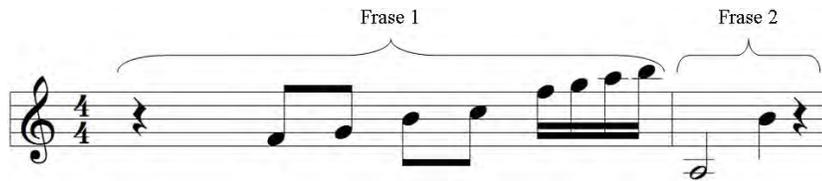
Puntos muertos (*impasses*): Si un punto muerto es declarado, el sistema decrementa en 10% el valor de la constante de resolución y de la constante ACAS, luego vuelve a *engagement*. El objetivo es hacer más fácil encontrar un átomo similar o igual. Si después de decrementar 10% las constantes el sistema es capaz de hacerlo, los valores de las constantes regresan a sus valores originales. Si después de modificar las constantes el sistema no puede encontrar ningún átomo, el sistema intenta una segunda estrategia: selecciona aleatoriamente una frase de su base de conocimiento, la añade a la melodía en progreso y regresa a *engagement*. El objetivo de esta estrategia es modificar

el contexto de forma que se pueda asociar algún átomo. Si el sistema falla otra vez un punto muerto es declarado como irrompible y el ciclo termina.

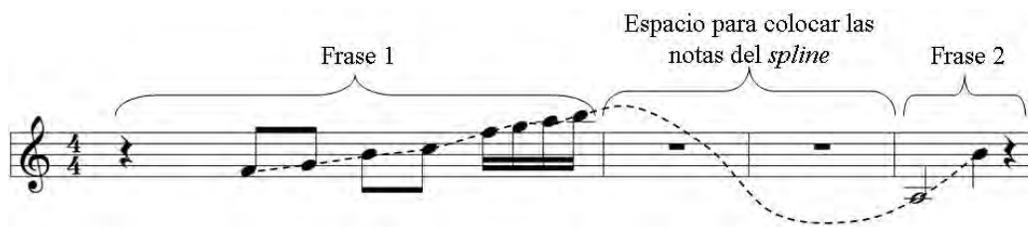
Coherencia: Para el sistema una melodía es coherente cuando no hay cambios abruptos de notas entre dos frases continuas. Por ejemplo, imaginemos que una melodía está formado por dos frases, y que la última nota de la primera frase es el primer intervalo y que la primera nota de la segunda frase es el séptimo intervalo. En este caso la distancia entre ellas es grande y por lo tanto la coherencia se ha roto. *La máxima distancia permitida entre notas continuas de dos frases distintas es un parámetro definido por el usuario del sistema.*

ERMEG cuenta con dos estrategias para resolver el problema de la coherencia. Una de ellas es aplicar una transposición proporcional a la distancia del salto que separa las frases. La transposición proporcional consiste en mover las notas de la frase en una cantidad de intervalos. *La proporción que utiliza el sistema es un parámetro del sistema que puede ser libremente modificado por el usuario.* La otra es construir un *spline* natural e interpolar notas en dos compases que unen las dos frases sin modificarlas. La construcción del *spline* consiste en generar una curva de interpolación. La curva generada debe colocar todos los puntos que ya estaban (i.e. las notas de las dos frases), y extenderse a través de dos compases intermedios generando una curva. Dado que se utiliza un *spline* natural, la curvatura es mínima, por lo tanto la transición de notas será suave.

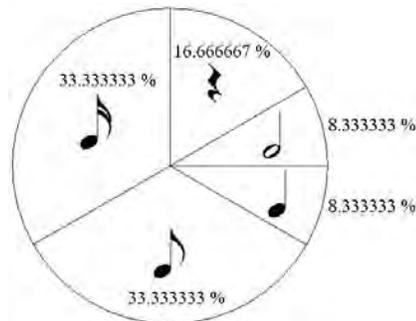
El número de notas interpoladas en el *spline* está determinado por sus duraciones. Cuando se generan las notas a interpolar se construye una cadena de notas con duración pero sin *sin altura* (tono). Por ejemplo, se genera una nota de tono X con duración de 1 tiempo. El *spline* abarca dos compases, así que se pueden insertar tantas notas como quepan en esa duración. Las duraciones que tendrá cada nota se seleccionan utilizando un mecanismo de ruleta donde unas duraciones tienen más probabilidad de ser seleccionadas que otras. Las probabilidades de la ruleta se forman con las duraciones de las frases vecinas al *spline*. En la figura 4.23 se muestra un ejemplo de dos frases que rebasan la máxima distancia permitida. En ella se encuentran varios tipos de duraciones. Existen dos silencios de 1 tiempo, cuatro corcheas, cuatro semicorcheas, una blanca y una negra. Con esos datos contruimos la ruleta de acuerdo a su frecuencia de aparición. La ruleta obtenida se muestra en la figura 4.24. Con ella se seleccionan las duraciones de la notas o silencios en el *spline*.



, **Figura 4.23:** Dos frases que rebasan la máxima distancia permitida. La distancia que separa las notas final de la frase 1 e inicial de la frase 2 es de 13 tonos



, **Figura 4.25:** Spline que coloca las notas de las frases 1 y 2 y que se dibuja entre ellas en dos compases.

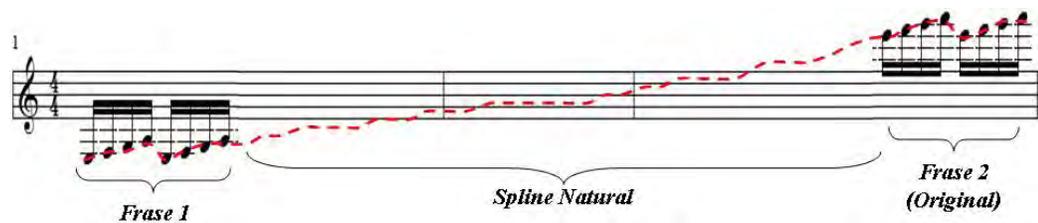


, **Figura 4.24:** Ruleta para seleccionar las duraciones de las notas en el spline

Conociendo el número de notas a interpolar y sus duraciones se construye una curva que toma en cuenta la posición de las notas (solo notas) en las frase 1 y 2 y que se extiende sobre dos compases. En la figura 4.25 se muestra la curva. Por último, sobre la curva se colocan las notas con su duración, la altura de la nota (tono) se determina por la posición de la curva en ese punto. En la figura 4.26 se muestra un ejemplo de dos frases que fueron puestas juntas por *engagement* y que rompen la coherencia de la melodía. En el



, **Figura 4.26:** Dos frases puestas juntas por engagement

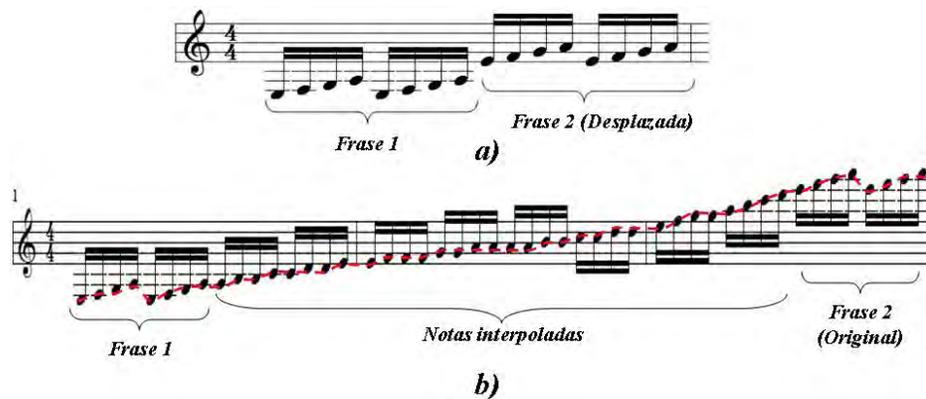


, **Figura 4.27:** Curva del Spline Natural

ejemplo de la figura la distancia que separa las notas en la unión de las frases es de $15\frac{1}{2}$ tonos. Para la transposición, la proporción que se utilizó fue de $\frac{3}{4}$. Con lo que, $\frac{3}{4}$ de $15\frac{1}{2}$ tonos es 11.625 tonos. Redondeando el valor tenemos una transposición de 12 tonos, es decir, que las notas de la segunda frase se acercan 12 tonos a la primera. Para construir el *spline*, en la figura 4.27 se muestra la curva generada que coloca los puntos y que se dibuja en los dos compases sin notas interpoladas.

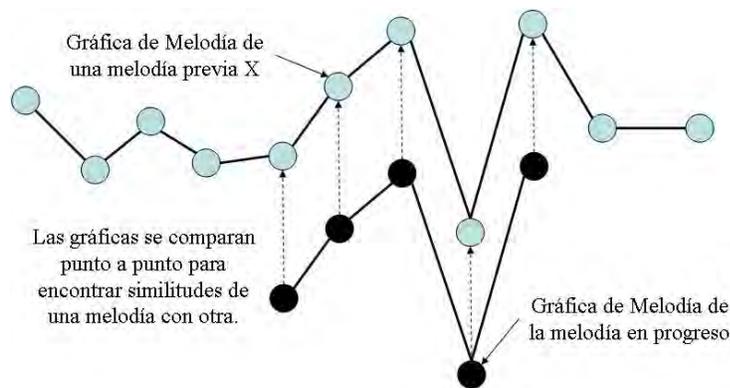
En la figura 4.28a se muestra la transposición de las frase 2 en 12 tonos. En b se muestra el *spline* generado y las notas interpoladas en dos compases que unen a las frases originales las cuales se ajustan a la curva.

Evaluación de la novedad: el sistema compara la melodía en progreso con todas las melodías previas. Si son similares el sistema decrementa el valor de las constantes de resolución y ACAS para encontrar átomos novedosos que asociar. La comparación que realiza *reflection* para determinar si la melodía en progreso es similar a las melodías previas, se lleva a cabo utilizando las gráficas de melodía y ritmo. Tales gráficas se obtuvieron durante la construcción de las estructuras de conocimiento y se mencionaron en la sección 4.3. Cuando el sistema compara la melodía en progreso con las melodías previas, el sistema busca encontrar que la melodía en progreso sea similar hasta un cierto porcentaje permitido. El porcentaje permitido es un parámetro modifi-



, **Figura 4.28:** *Modificaciones hechas por reflection. En a) Reflection aplicó una transposición de 12 tonos a la frase 2. En b) Se ha construido un spline natural y se han interpolado notas en dos compases.*

cable por el usuario. Si lo rebasa se considera que está copiando alguna de las melodías previas. La comparación se realiza en todos los niveles de la gráficas de melodía y ritmo, es decir, se comparan punto a punto las gráficas. La mayor similitud se guarda y al final esa es la que se compara con el parámetro de similitud permitido. En la figura 4.29 se muestra un ejemplo de comparación de melodías.



, **Figura 4.29:** *Las gráficas se comparan punto a punto para encontrar similitudes de una melodía con otra. En este caso la gráfica de la melodía en progreso es 100 % similar a un fragmento de la melodía previa X. El sistema está copiando.*

El ciclo E-R termina cuando un punto muerto no se puede romper, o después del número de ciclos completos del E-R que el usuario defina.

4.5. Resumen de los parámetros del sistema.

El sistema cuenta con 10 parámetros los cuales ayudan al usuario manejar el sistema. Estos parámetros son muy importantes ya que ERMEG es un instrumento de investigación. Con ellos, los investigadores tienen mucha libertad para experimentar con el sistema.

Los parámetros del sistema se muestran en la tabla 4.11. En ella encontramos el nombre de ellos, su descripción y una referencia a la sección donde se describen.

, **Tabla 4.11:** Resumen de parámetros del sistema.

Parámetro	Descripción	Referencia
Ciclos ER	Este parámetro sirve para controlar el número de ejecuciones completas del ciclo <i>Engagement-Reflection</i> .	Ver sección 4.4 página 76
Ciclos E	Este parámetro sirve para controlar el número de ejecuciones del ciclo <i>Engagement</i> .	Ver sección 4.4.1 página 76
Constante de Resolución Local	Con este parámetro se especifica la constante de resolución para los valores locales.	Ver página 80
Constante de resolución Global	Con este parámetro se especifica la constante de resolución para los valores globales.	Ver página 80
ACAS Superior	Este valor indica el porcentaje superior del rango de similitud que usa el sistema para encontrar átomos similares.	Ver página 80
Continúa en la siguiente página		

Tabla 4.11 – Continúa de la página anterior

Parámetro	Descripción	Referencia
ACAS Inferior	Este valor indica el porcentaje inferior del rango de similitud que usa el sistema para encontrar átomos similares.	Ver página 80
Máxima Diferencia entre notas	Es la distancia máxima que puede existir entre dos notas consecutivas de distintas frases.	Ver página 84
Máxima Copia	Es el porcentaje máximo de copia que puede tener la melodía en composición con las melodías previas.	Ver página 86
Proporción del salto	Es la proporción en la que se efectúa una transposición.	Ver página 84
Tamaño de las frases	Con este valor se especifica el tamaño de las subdivisiones de frases.	Ver sección 4.3.1

4.6. Discusión

ERMEG es un sistema que se basa en una representación computacional de emociones y tensiones y en el modelo computacional *Engagement-Reflection* para componer melodías musicales. Con este sistema se ha tratado de mostrar cómo nos pueden servir las representaciones computacionales de emoción y tensión para guiar una composición en el modelo E-R.

En el siguiente capítulo se muestra un ejemplo paso a paso de la ejecución del sistema para que toda la descripción que se ha visto en este capítulo se complemente.

Capítulo 5

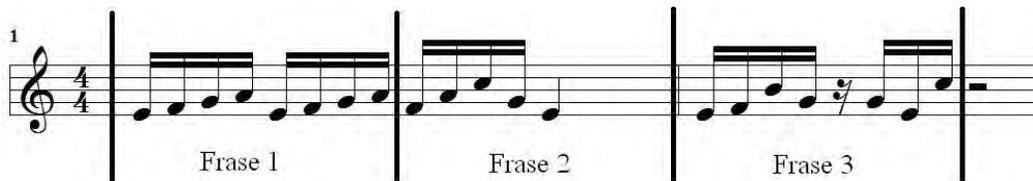
Cómo trabaja ERMEG

El objetivo de este capítulo es mostrar un ejemplo completo de ejecución de ERMEG. Está dividido en dos partes que corresponden a los procesos principales que lleva a cabo el sistema. Con esto se completa la descripción de la estructura del programa vista en el capítulo 4.

- En la primera parte se aborda la construcción de las estructuras de conocimiento que se utilizan para componer una pieza musical. Estas estructuras se obtienen de un conjunto de melodías previas que son suministradas al sistema por el usuario. De ellas se extrae información relevante para el proceso de composición.
- En la segunda parte se describe el funcionamiento del ciclo *engagement-reflection* para componer una melodía musical. En él se utilizan las estructuras de conocimiento obtenidas. Al final de la ejecución del ciclo se obtiene una melodía nueva.
- Al término del capítulo se muestran algunos ejemplos de composiciones realizadas por el sistema.



, *Figura 5.1: Ejemplo de una melodía previa.*



, *Figura 5.2: Melodía previa de ejemplo de la figura 5.1 dividida en frases*

5.1. Construcción de estructuras de conocimiento.

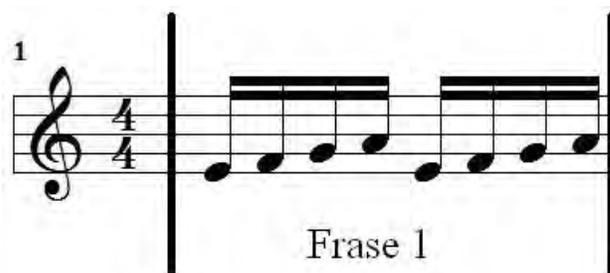
Como se mencionó en la sección 4.3.1 las melodías se deben dividir para poder analizarlas. Con el ejemplo de la figura 5.1 se muestra como el sistema divide las melodías previas. Los datos obtenidos se utilizan para generar gráficas de tensión, emoción, melodía y ritmo. Las estructuras de conocimiento mencionadas en las secciones 4.3.2 y 4.3.3 contextos y átomos, se construyen empleando la información de tales gráficas.

Las frases pueden ser de diferentes duraciones. En este sencillo ejemplo la melodía se subdivide en frases de dos tiempos. El resultado se muestra en la figura 5.2.

En el capítulo 4 se mencionó que a las frases composicionales se les calculan valores numéricos de características intrínsecas. Esto se detalla a continuación.

5.1.1. Características de las frases

Para mostrar el modo de operación del sistema cuando evalúa las características de las melodías utilizamos la primera frase del ejemplo de la figura 5.2. Con las otras se procede de la misma forma. En la figura 5.3 tenemos la frase a la que le calcularemos la velocidad, variación y diferencia interválica.



, *Figura 5.3: Frase 1 de la melodía de muestra*

Para calcular la velocidad utilizamos el número de notas y la duración de la frase (ver sección 3.3, página 46: Velocidad). La relación que exista entre estos valores nos indica la densidad de notas que tiene esa frase. Ese valor representa la velocidad de la ejecución.

En la primera frase, el número de notas es ocho, y la duración es de dos tiempos. Por lo tanto tenemos:

$$Velocidad = \frac{8}{2} \Rightarrow Velocidad = 4$$

Para calcular la diferencia interválica utilizamos la distancia que existe entre la primera y la última nota de la frase. Este ejemplo inicia con una nota Mi (E) y termina con una nota La (A), por lo tanto esta frase tiene una diferencia interválica de $2\frac{1}{2}$ tonos:

$$DiferenciaInterválica = 2.5$$

Para calcular la variación interválica utilizamos las distancias que hay de una nota a la siguiente y obtenemos el promedio de las distancias en toda la frase. El valor obtenido es la variación interválica. En este ejemplo la distancia; entre la primera y segunda notas es de $\frac{1}{2}$ tono; entre la segunda y la tercera es de 1 tono; entre la tercera y la cuarta es de 1 tono; entre la cuarta y la quinta es de $2\frac{1}{2}$ tonos; entre la quinta y sexta es de $\frac{1}{2}$ tono; entre la sexta y séptima es de 1 tono y entre la séptima y octava es de 1 tono. Por lo tanto tenemos que:

$$VariaciónInterválica = \frac{0.5+1+1+2.5+0.5+1+1}{7}$$

$$VariaciónInterválica = 1.071..$$

5.1.2. Tensión y Emoción

Los valores calculados se emplean como entrada del sistema difuso visto en la sección 4.2. El sistema devuelve los valores numéricos de tensión y emoción de la frase.

Lo primero que se hace en el sistema difuso es *fuzzyficar* las variables de entrada. Con este proceso obtenemos variables lingüísticas. Éstas, van acompañadas por predicados que corresponden a los conjuntos difusos. Cada variable de entrada obtiene diferentes valores de membresía a cada conjunto difuso. A continuación se muestra cómo se asignan los predicados y cómo se obtienen los valores de membresía. Tomamos como entrada del sistema difuso los valores de las características de la frase composicional de la figura 5.3.

Los valores calculados de las características de la frase musical deben ubicarse en el universo de discurso de las variables lingüísticas. Luego se proyectan perpendicularmente al eje de los valores del universo de discurso. Se encuentran los puntos en donde las proyecciones cortan las rectas de los conjuntos difusos. La altura a la que se corten las rectas de los conjuntos difusos corresponde al valor de membresía que tiene un conjunto difuso para un valor particular en el universo de discurso.

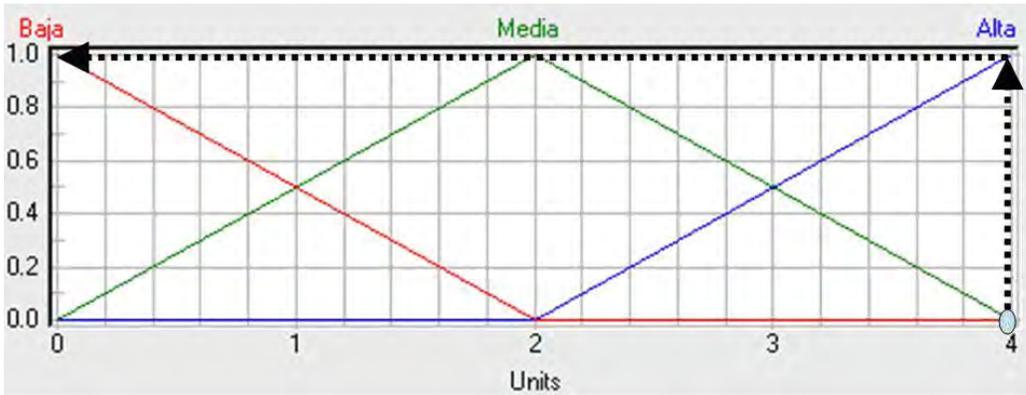
Para la variable velocidad el valor obtenido fue 4. En la figura 5.4 se muestran los conjuntos difusos de la variable velocidad, la ubicación del valor obtenido y el valor de membresía al conjunto difuso de velocidad alta, el cual en este caso es de 1.0.

Lo mismo sucede para las variables diferencia y variación interválica. Los valores obtenidos fueron en estos casos: 2.5 y 1.071 respectivamente. En las figuras 5.5 y 5.6 se muestran también sus funciones de membresía, los conjuntos difusos a los que pertenece estos valores y las membresías correspondientes a cada conjunto difuso.

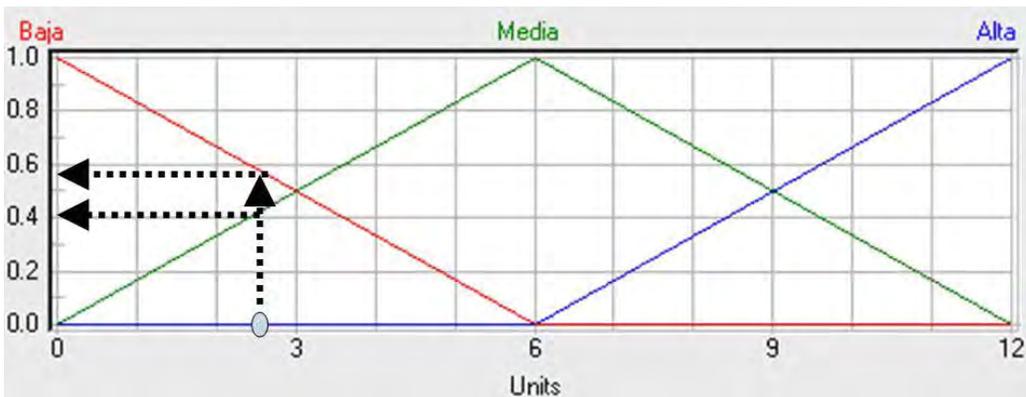
El resultado que se obtiene al fuzzyficar el valor de 2.5 de diferencia interválica, es que este valor pertenece a dos conjuntos difusos: el de diferencia baja y el de diferencia media. Los valores de membresía con los que pertenece son 0.583 y 0.416 respectivamente.

Cuando se fuzzifica el valor 1.071 de variación interválica se observa que este valor pertenece a dos conjuntos difusos: el de variación media y el de variación alta. Los valores de membresía con los que pertenece a cada uno son 0.971 y 0.028 respectivamente.

Una vez que hemos fuzzyficado los valores de las características de las



, **Figura 5.4:** Funciones de membresía de “Velocidad”, ubicación del valor 4 de velocidad, su membresía al conjunto difuso de velocidad alta y su correspondiente valor de membresía.



, **Figura 5.5:** Funciones de membresía de “Diferencia”, ubicación del valor 2.5 de diferencia. Membresía a los conjuntos difusos diferencia baja y diferencia media y sus correspondientes valores de membresía.



, **Figura 5.6:** Funciones de membresía de “Variación”, ubicación del valor 1.071 de variación. Membresía a los conjuntos difusos variación baja y variación media y sus correspondientes valores de membresía.

frases musicales y que hemos obtenido las correspondientes variables lingüísticas con sus correspondientes valores de membresía tenemos que aplicar un conjunto de reglas difusas para obtener los valores de tensión y emoción de la frase. Este conjunto de reglas difusas ha sido generado utilizando resultados de investigaciones de música y emociones. [Juslin & Sloboda, 2001], [Dalla Bella, Peretz, Rousseau & Gosselin, 2001], [Khalfa, Roy, Rainville, Dalla Bella, & Peretz, 2008].

Hasta el momento se tiene:

Velocidad_Alta	con valor de membresía 1.0
Diferencia_Baja	con valor de membresía 0.583
Diferencia_Media	con valor de membresía 0.416
Variación_Media	con valor de membresía 0.971
Variación_Alta	con valor de membresía 0.028

Con esto en mente debemos seleccionar del conjunto de reglas difusas, visto en la sección 4.2.2, aquellas reglas que incluyen combinaciones de estas variables. Las reglas seleccionadas son las siguientes:

Para emoción:

1. IF Velocidad = Alta AND Diferencia = Baja AND Variación = Media THEN Emoción = Alegre

Reglas seleccionadas	Variables lingüísticas	Valores de membresía
Regla 1	Alegre	0.583
Regla 2	Muy_Alegre	0.416
Regla 3	Muy_Alegre	0.028
Regla 4	Muy_Alegre	0.028
Regla 5	Calmado	0.583
Regla 6	Neutral	0.416
Regla 7	Neutral	0.028
Regla 8	Ansioso	0.028

, **Tabla 5.1:** Valores obtenidos aplicando las reglas de membresía.

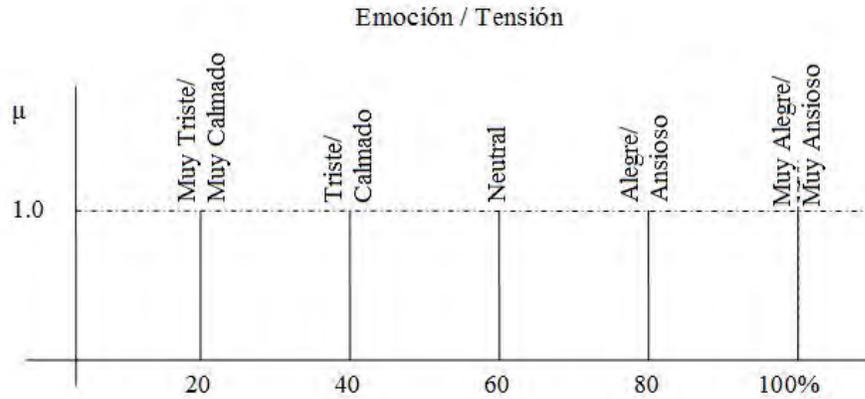
2. IF Velocidad = Alta AND Diferencia = Baja AND Variación = Alta
THEN Emoción = Muy_Alegre
3. IF Velocidad = Alta AND Diferencia = Media AND Variación = Media
THEN Emoción = Muy_Alegre
4. IF Velocidad = Alta AND Diferencia = Media AND Variación = Alta
THEN Emoción = Muy_Alegre

Para tensión

5. IF Velocidad = Alta AND Diferencia = Baja AND Variación = Media
THEN Tensión = Calmado
6. IF Velocidad = Alta AND Diferencia = Baja AND Variación = Alta
THEN Tensión = Neutral
7. IF Velocidad = Alta AND Diferencia = Media AND Variación = Media
THEN Tensión = Neutral
8. IF Velocidad = Alta AND Diferencia = Media AND Variación = Alta
THEN Tensión = Ansioso

Luego, empleando los valores de membresía obtenidos, aplicamos las reglas seleccionadas. El resultado se muestra en la tabla 5.1.

Con los valores obtenidos para cada variable lingüística aplicamos el proceso de desfuzzificación para obtener los valores de tensión y emoción. Para hacer



, *Figura 5.7: Universo de discurso de las variables Tensión y Emoción y conjuntos difusos mostrados como singletons*

esto empleamos los valores diáfanos, o valores en los que las variables lingüísticas de salida (consecuente) tienen valor de membresía de 1.0, es decir, que tratamos a las variables del consecuente como *singletons*¹. Esto para hacer más fácil y rápido el cálculo en el sistema. La figura 5.7 muestra el universo de discurso de salida de las variables de tensión y emoción con *singletons*. Para la desfuzzificación empleamos el método de centro de área (COA), y obtenemos expresiones como las siguientes²:

$$Emoción = \frac{(Alegre*0.583)+(Muy_Alegre*0.416)+(Muy_Alegre*0.028)+(Muy_Alegre*0.028)}{0.583+0.416+0.028+0.028}$$

$$Emoción = \frac{(80*0.583)+(100*0.416)+(100*0.028)+(100*0.028)}{0.583+0.416+0.028+0.028}$$

$$Emoción = 88.96$$

$$Tensión = \frac{(Calmado*0.583)+(Neutral*0.416)+(Neutral*0.028)+(Ansioso*0.028)}{0.583+0.416+0.028+0.028}$$

¹Son conjuntos difusos representados por un solo punto en el universo de discurso con membresía de 1.0

²En estas expresiones sustituimos las variables lingüísticas y utilizamos en su lugar el valor de los *singletons* para esa variable.

	Tensión	Emoción
Frase 1	247.50	88.96
Frase 2	208.94	71.818
Frase 3	279.64	82.50

, **Tabla 5.2:** Valores de tensión y emoción para las tres frases de la melodía previa del ejemplo

$$Tensión = \frac{(40*0.583)+(60*0.416)+(60*0.028)+(80*0.028)}{0.583+0.416+0.028+0.028}$$

$$Tensión = 49.50$$

Este valor de tensión es además afectado por la octava promedio de la frase, pues como se mencionó mientras mayor sea la octava, mayor tensión genera; y mientras menor sea la octava, menor tensión genera. La octava promedio de la frase de ejemplo es 5, por lo tanto tenemos que:

$$Tensión = 49.50 * 5 \Rightarrow Tensión = 247.50$$

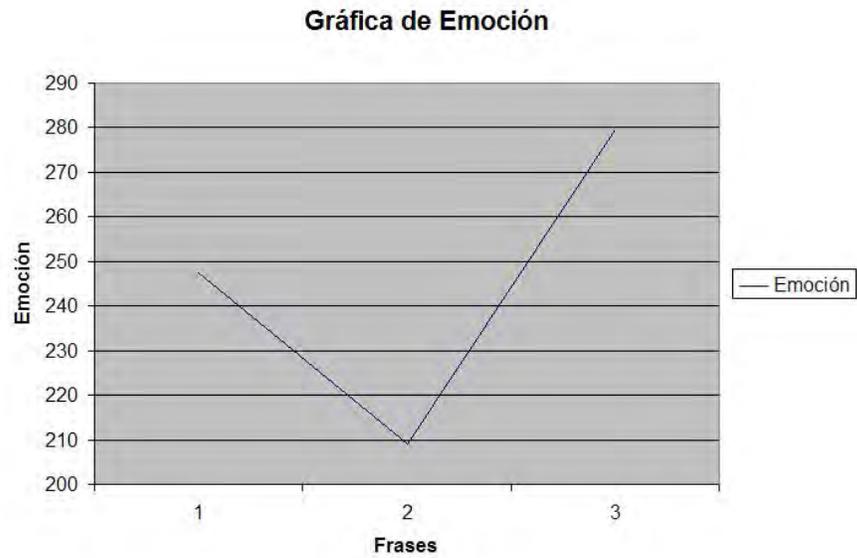
La misma operación se repite para las demás frases. Al final obtenemos, para las tres frases del ejemplo, los valores de tensión y emoción que se muestran en la tabla 5.2. Las gráficas se muestran en las figuras 5.8 y 5.9.

De esta forma se construyen las gráficas de tensión y emoción de las melodías previas. Este proceso se sigue para todas las frases de todas las melodías previas. En este ejemplo se utilizó una melodía de tres frases. Una melodía previa usualmente es más larga.

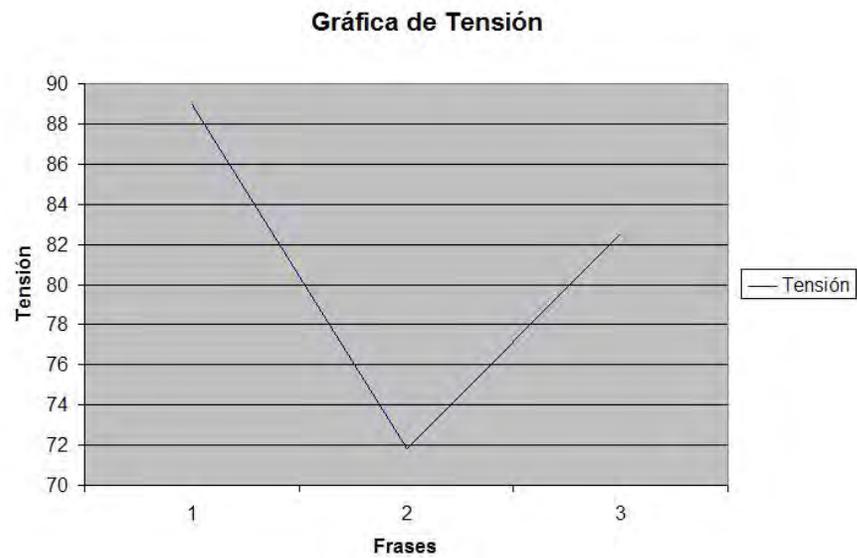
Con las gráficas podemos obtener rápidamente el vector de tensiones y emociones históricas para una frase particular. Esto nos es útil al momento de construir los contextos y átomos que el sistema también utiliza para componer.

5.1.3. Melodía y Ritmo

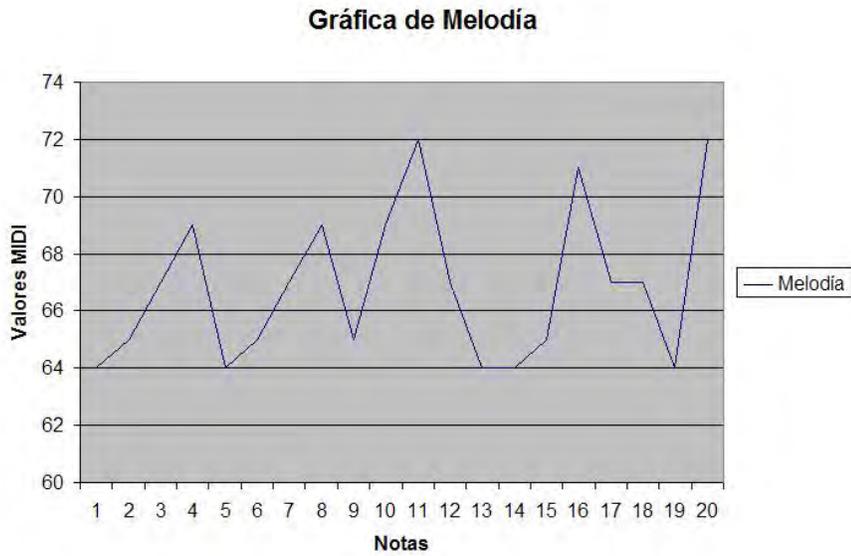
Dos estructuras más que genera el sistema son las gráficas de melodía y ritmo. Éstas se utilizan durante *reflection* para comprobar que ERMEG no está copiando las melodías previas. La forma de construir estas gráficas se mostró en la sección 4.3.1, y las gráficas se muestran en las figuras 5.10 y 5.11. Igual que en el caso anterior este proceso se sigue para todas las frases de todas las melodías previas.



, *Figura 5.8: Gráfica de Emoción obtenida de las tres frases de la melodía previa de ejemplo.*



, *Figura 5.9: Gráfica de Tensión obtenida de las tres frases de la melodía previa de ejemplo.*



, *Figura 5.10: Gráfica de Melodía obtenida de la melodía previa de ejemplo.*



, *Figura 5.11: Gráfica de Ritmo obtenida de la melodía previa de ejemplo.*

Contexto 1	Contexto 2	Contexto 3
<u>Local</u>	<u>Local</u>	<u>Local</u>
Emoción: 88.96	Emoción: 71.818	Emoción: 82.50
Tensión: 247.50	Tensión: 208.94	Tensión: 279.64
<u>Históricos</u>	<u>Históricos</u>	<u>Históricos</u>
Emoción:[]	Emoción:[88.96]	Emoción:[88.96, 71.818]
Tensión:[]	Tensión:[247.50]	Tensión:[247.50, 208.94]

, **Figura 5.12:** Contextos de la melodía previa de ejemplo.

5.1.4. Contextos y átomos

Una vez que tenemos las gráficas de tensión y emoción de las melodías previas podemos construir los contextos y los átomos; dos estructuras de mucha importancia en el sistema.

Contextos

Los contextos, como ya se mencionó en la sección 4.3.2, se construyen con los valores de tensión local, tensión global o vector histórico de tensiones, emoción local y emoción global o vector histórico de emociones.

Los contextos de las tres frases de la melodía previa de ejemplo y se muestran en la figura 5.12.

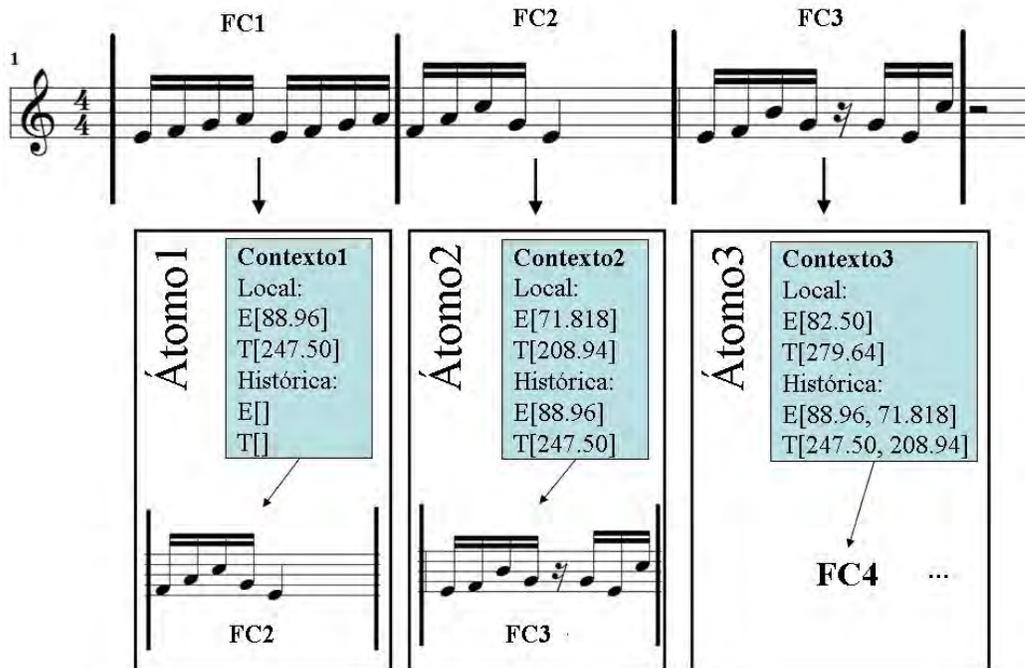
Átomos

En la sección 4.3.3 se mostró la forma de construir los átomos. En la figura 5.13 se muestran los átomos obtenidos para las frases de la melodía previa de ejemplo. Para aclarar con un ejemplo, el Contexto 2 pertenece a la frase 2 de la melodía previa. Para crear el átomo asociamos la frase 3 al Contexto 2. *La frase 3 es la opción lógica de ser ejecutada cuando se encuentra un contexto como el contexto2.*

5.2. Ciclo *engagement-reflection*.

Con las gráficas de tensión, emoción, melodía y ritmo, los contextos y átomos, el sistema puede iniciar un ciclo *engagement-reflection* para componer una melodía nueva. la cantidad de repeticiones puede ser modificado por el

usuario. Este parámetro es llamado CiclosER, en este caso su valor será de 3.

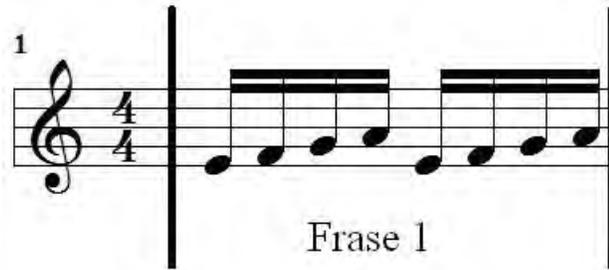


, *Figura 5.13: Los átomos de la melodía previa de ejemplo. En ellos podemos encontrar la opción lógica a ser ejecutada.*

Para mostrar el funcionamiento del sistema en el ciclo E-R, se genera una melodía comenzando con una frase inicial propuesta, y se utiliza la base de conocimiento del sistema que consta de 19 melodías previas que tienen un promedio de 37.78 frases cada una, y que para este ejemplo tienen una longitud de dos tiempos. La base de conocimiento consta de 699 átomos.

Cuando el sistema ejecuta el ciclo *engagement-reflection*, este inicia en el estado *engagement*. En este estado construimos una cadena de ideas musicales asociadas. En *reflection* se resuelven incongruencias musicales (grandes saltos en la melodía) y se verifica que no se estén copiando melodías previas. Si *engagement* no encuentra ideas que asociar declara un punto muerto (*impass*) y es trabajo de *reflection* tratar de romperlo.

En la sección 4.4.1 se describió la forma en que trabaja el ciclo *engagement*. La primer condición para empezar trabajar, es obtener una frase inicial del



, **Figura 5.14:** Frase 1 de la melodía de ejemplo que utilizaremos como frase inicial.

Contexto Inicial
<u>Local</u>
Emoción: 88.96
Tensión: 247.50
<u>Históricos</u>
Emoción: []
Tensión: []

, **Tabla 5.3:** Contexto Inicial

usuario. Por claridad del ejemplo, vamos a suponer que el usuario proporciona como frase inicial la misma que analizamos al inicio de este capítulo (ver figura 5.14). El sistema calcula los valores de tensión y emoción y actualiza el Contexto, este es el contexto inicial de la melodía en progreso. Este es empleado para localizar en memoria átomos que sean iguales o similares y se obtendrán las FC asociadas a tales átomos. De las FC encontradas el sistema selecciona una de ellas para agregarla a la melodía en progreso como la siguiente acción musical a ser ejecutada, con esto último el contexto cambia, se actualiza y el proceso se repite.

Los niveles de tensión y emoción ya habían sido calculados previamente y son los siguientes:

$$Tensión = 247.50 \text{ Y } Emoción = 88.96$$

Se actualiza el contexto (inicial) y queda como se ve en la tabla 5.3: La ejecución de el ciclo *engagement* se detalla a continuación.

5.2.1. Inicia Ciclo *Engagement*

El ciclo *Engagement* tendrá 3 repeticiones; este parámetro es ajustado por el usuario.

Los valores de las constantes de resolución y ACAS (ver página 80) son:

Resolución Local	90.0 %
Resolución Global	70.0 %
ACAS Superior	100.0 %
ACAS Inferior	50.0 %

Primer ciclo *Engagement*

Los valores de las características de la frase actual (inicial) son:

Velocidad	4.0
Diferencia	2.5
Variación	1.0714285714285714

Los valores de tensión y emoción calculados son:

Tensión	247.52252252252256
Emoción	88.96396396396395

Teniendo en mente los valores de las constantes de resolución y ACAS usamos los contextos como muestra para buscar en memoria átomos que sean similares al contexto actual. La forma de hacer la comparación (*match*) se mostró en la página 82 (sección 4.4.1). Como resultado obtenemos los átomos con sus correspondientes valores de similitud que se muestran en la tabla 5.4.

Cantidad total de Átomos similares o iguales encontrados: 17

El átomo seleccionado aleatoriamente fue: Átomo365

La melodía hasta el momento se muestra en la figura 5.15.

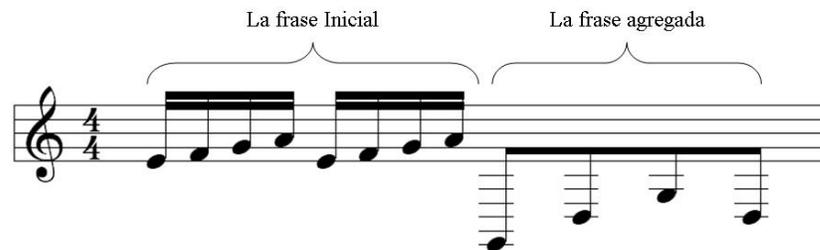
Segundo ciclo *Engagement*

Los valores de las características de la frase actual son:

Velocidad	2.0
Diferencia	3.5
Variación	2.8333333333333335

Átomos	Porcentaje de Similitud
Atomo75	50.0 %
Atomo257	50.0 %
Atomo258	62.99605249474365 %
Atomo281	50.0 %
Atomo282	62.99605249474365 %
Atomo359	50.0 %
Atomo365	100.0 %
Atomo366	50.0 %
Atomo395	100.0 %
Atomo396	62.99605249474365 %
Atomo430	62.99605249474365 %
Atomo471	100.0 %
Atomo472	62.99605249474365 %
Atomo494	50.0 %
Atomo509	50.0 %
Atomo603	100.0 %
Atomo604	62.99605249474365 %

, **Tabla 5.4:** 17 átomos seleccionados por el sistema en la primera ejecución de engagement, pues su porcentaje de similitud con el contexto actual de la melodía está dentro del rango de la constante ACAS.



, **Figura 5.15:** La Melodía en progreso después de la primera ejecución de engagement

Los valores de tensión y emoción calculados son:

Tensión	178.5326086956522
Emoción	59.99999999999999

Nuevamente considerando los valores de las constantes de resolución y ACAS usamos los contextos como muestra para buscar en memoria átomos que sean similares al contexto actual. La forma de hacer la comparación (*match*) se mostró en la página 82. Como resultado obtenemos los átomos con sus correspondientes valores de similitud que se muestran en la tabla 5.5.

, **Tabla 5.5:** 37 átomos seleccionados por el sistema en la segunda ejecución de engagement, pues su porcentaje de similitud con el contexto actual de la melodía está dentro del rango de la constante ACAS.

Átomos	Porcentaje de Similitud
Atomo75	62.99605249474365 %
Atomo76	81.49802624737183 %
Atomo77	64.43034031735094 %
Atomo78	54.18323378305069 %
Atomo258	81.49802624737183 %
Atomo259	56.74820211933147 %
Atomo281	62.99605249474365 %
Atomo283	56.74820211933147 %
Atomo323	54.18323378305069 %
Atomo324	60.312457149764384 %
Atomo360	50.0 %
Atomo365	62.99605249474365 %
Atomo366	100.0 %
Atomo367	66.35535464043733 %
Atomo368	68.6814412668964 %
Atomo369	52.75258795338228 %
Atomo395	62.99605249474365 %
Atomo396	68.50197375262817 %
Atomo397	56.74820211933147 %
Atomo398	68.6814412668964 %

Continúa en la siguiente página

Tabla 5.5 – Continúa de la página anterior

Átomos	Porcentaje de Similitud
Atomo399	52.75258795338228 %
Atomo429	62.99605249474365 %
Atomo430	62.99605249474365 %
Atomo431	80.78569495778827 %
Atomo433	60.312457149764384 %
Atomo471	62.99605249474365 %
Atomo472	100.0 %
Atomo473	51.92501432308638 %
Atomo474	68.6814412668964 %
Atomo482	62.99605249474365 %
Atomo483	64.43034031735094 %
Atomo494	62.99605249474365 %
Atomo495	81.49802624737183 %
Atomo496	50.0 %
Atomo603	50.0 %
Atomo604	100.0 %
Atomo605	66.35535464043733 %

Cantidad total de Átomos similares o iguales encontrados: 37

El átomo seleccionado aleatoriamente fue: Átomo75

La melodía hasta el momento se muestran en la figura 5.16.

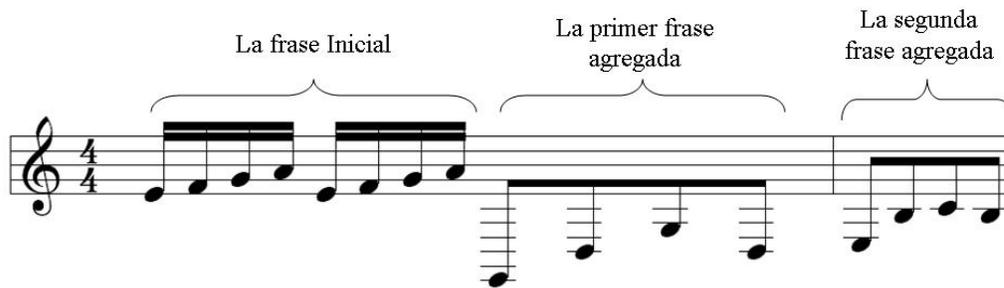
Tercer ciclo *Engagement*

Los valores de las características de la frase actual son:

Velocidad	2.0
Diferencia	3.5
Variación	1.5

Los valores de tensión y emoción calculados son:

Tensión	192.26190476190477
Emoción	60.0



, **Figura 5.16:** La Melodía en progreso después de la segunda ejecución de engagement. La pequeña barra en el pentagrama surge como consecuencia de que finalizó el primer compás e indica el principio del segundo.

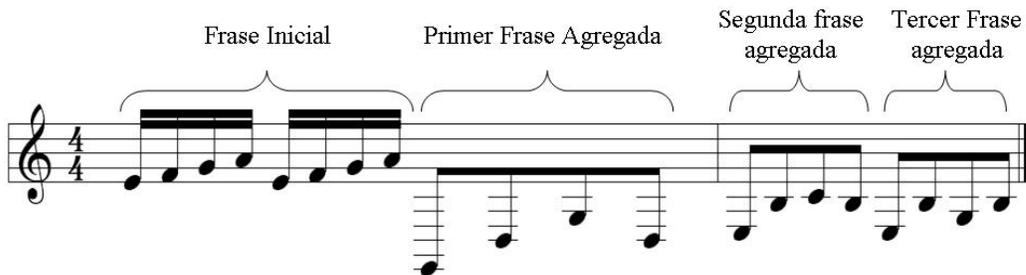
Considerando los valores de las constantes de resolución y ACAS usamos los contextos como muestra para buscar en memoria átomos que sean similares al contexto actual. La forma de hacer la comparación (*match*) se mostró en la página 82. Se obtuvieron los átomos y sus correspondientes valores de similitud que se muestran en la tabla 5.6.

, **Tabla 5.6:** 53 átomos seleccionados por el sistema en la tercera ejecución de engagement, pues su porcentaje de similitud con el contexto actual de la melodía está dentro del rango de la constante ACAS.

Átomos	Porcentaje de Similitud
Atomo76	80.78569495778827 %
Atomo77	90.39284747889414 %
Atomo78	78.54616911694792 %
Atomo79	69.85234742014926 %
Atomo80	63.22969308799982 %
Atomo81	57.99998388716773 %
Atomo82	53.74909973054214 %
Atomo258	64.43034031735094 %
Atomo259	66.35535464043734 %
Atomo260	59.86472785005154 %
Continúa en la siguiente página	

Tabla 5.6 – Continúa de la página anterior

Átomos	Porcentaje de Similitud
Atomo263	55.65874550280658 %
Atomo282	80.78569495778827 %
Atomo284	54.07017633355127 %
Atomo322	64.43034031735094 %
Atomo323	68.6814412668964 %
Atomo324	69.85234742014926 %
Atomo325	72.33196571510956 %
Atomo326	59.61433702420793 %
Atomo327	53.74909973054214 %
Atomo361	50.0 %
Atomo366	80.78569495778827 %
Atomo367	100.0 %
Atomo368	78.54616911694792 %
Atomo369	79.39223769053412 %
Atomo370	63.83999586794737 %
Atomo371	50.951030759832726 %
Atomo375	51.87314514310193 %
Atomo387	62.0025082170944 %
Atomo396	80.78569495778827 %
Atomo397	59.60715252110587 %
Atomo398	73.91268948315376 %
Atomo399	62.29247822376715 %
Atomo400	51.91895724443684 %
Atomo404	54.75778807854069 %
Atomo430	80.78569495778827 %
Atomo431	80.78569495778827 %
Atomo432	88.41089696699946 %
Atomo433	60.312457149764384 %
Atomo434	72.33196571510956 %
Atomo435	52.56538389687293 %
Atomo436	51.754662595099454 %
Atomo472	80.78569495778827 %
Atomo473	51.92501432308638 %
Atomo474	59.41448199930805 %
Continúa en la siguiente página	



, **Figura 5.17:** La Melodía en progreso después de la tercera ejecución de *engagement*

Tabla 5.6 – Continúa de la página anterior

Átomos	Porcentaje de Similitud
Atomo483	56.74820211933147 %
Atomo495	80.78569495778827 %
Atomo496	90.39284747889414 %
Atomo497	58.70365596734543 %
Atomo501	62.0025082170944 %
Atomo502	52.88260620826337 %
Atomo578	55.999987915375804 %
Atomo604	66.35535464043733 %
Atomo605	75.96250716154319 %

Cantidad total de Átomos similares o iguales encontrados: 53

El átomo seleccionado aleatoriamente fue: Átomo79

La melodía hasta el momento se muestran en la figura 5.17.

De esta manera finaliza el primer ciclo *engagement*. El control pasa ahora al estado *reflection*. Durante *reflection* el sistema revisa y evalúa la novedad y coherencia del material generado hasta el momento, y si fuera necesario lo modifica transponiendo o insertando notas. El detalle de cómo funciona el estado *reflection*, como evalúa la novedad, coherencia y como modifica el material generado se describió en la sección 4.4.2.

5.2.2. Inicia *Reflection*

Con el material generado durante *engagement*, que se muestra en la figura 5.17, continuamos en *reflection*. La ejecución va como sigue:

Los valores de los parámetros modificables por el usuario

Porcentaje máximo de copia permitido:	50.0 %	ver página 86
El máximo intervalo permitido es:	6.0 Tonos %	ver página 84

Los valores de copia en la melodía encontrados por el sistema fueron:

Porcentaje de copia de melodía	18.75 %
Porcentaje de copia del ritmo	24.242424242424242 %

Lo primero que hace el sistema es verificar que la melodía en progreso no sea una copia de alguna de las melodías previas. Para ello compara cada una de ellas como se explicó en la sección 4.4.2. Para este ejemplo, la constante de similitud es igual al 50 %. Dado que el máximo valor entre la melodía en progreso y la melodía más similar es del 24.242 %, el sistema clasifica la melodía en progreso como novedosa. Como no resultó en copia de las melodías previas, el sistema no tiene que modificar los parámetros de *engagement* (ACAS, Constante de resolución) para futuras ejecuciones.

Cuando el sistema pasa a la revisión de la coherencia de la melodía comprueba que no existan saltos grandes entre dos frases puestas juntas. Un salto grande corresponde con uno que sea mayor que el máximo permitido por el usuario. En este caso se permite que los saltos entre frases sean como máximo de 6 tonos.

Con fines explicativos dejaremos de nombrar a la frase inicial como tal y la llamaremos frase 1. Las siguientes frases se enumeraran consecutivamente, por lo tanto, ahora tenemos una melodía con el siguiente conjunto de frases: frase 1, frase 2, frase 3 y frase 4.

Entre nota final de la frase 1 y la primer nota de la frase 2 la distancia que existe es de 13 tonos. Con esto se excede con mucho la máxima distancia permitida y hay que modificar para que la melodía sea coherente. Como se mencionó en la sección 4.4.2 el sistema emplea dos mecanismos para resolver estas inconsistencias musicales y los selecciona de forma aleatoria. Para este primer caso la selección corresponde con una transposición, y esta es de 9 Tonos. El cálculo de los tonos de la transposición es de aproximadamente $\frac{2}{3}$ de la distancia total que separa a las frases. Para estas dos frases la distancia total es de 13 y $\frac{2}{3}$ de 13 son 8.66666, redondeando, la transposición para acercarlas es de 9 tonos. La primera nota de la segunda frase es Sol de la tercer octava (G3); esa misma nota 9 tonos más arriba es Re bemol de la quinta octava (Db5), de esta forma se ha movido. El mismo procedimiento se repite con todas las notas de la segunda frase.

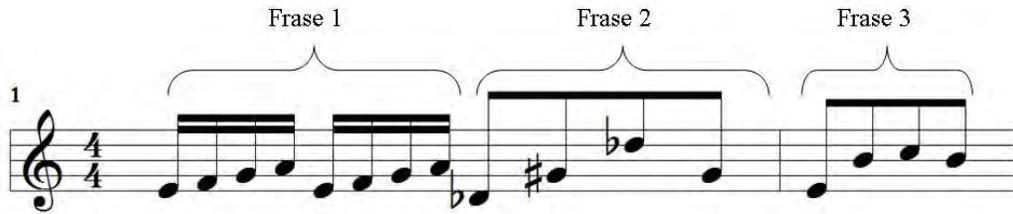
Figure 5.18 consists of two musical staves, labeled a) and b). Both staves are in 4/4 time and start with a treble clef and a '1' above the staff. Staff a) shows a melody with two phrases. The first phrase, labeled 'Frase Inicial', consists of two measures of eighth notes: G4-A4-B4, C5-B4-A4, G4. The second phrase, labeled 'Primer Frase agregada por engagement', consists of two measures of eighth notes: G4-A4-B4, C5-B4-A4, G4. Staff b) shows the same first phrase, but the second phrase is modified. It starts with a half note G4, followed by a quarter note A4 with a sharp sign (#), and a quarter note B4 with a flat sign (b). This modification is intended to bring the second phrase closer to the first.

, **Figura 5.18:** Modificación hecha por reflection entre la primer y segunda frase de la melodía en progreso. En a) se observa la figura original generada en los ciclos de engagement, en b) se observa la modificación hecha para acercar la segunda frase a la primera.

Figure 5.19 shows two musical phrases, labeled 'Frase 2' and 'Frase 3', on a single staff. Frase 2 consists of four notes: G4 with a flat sign (b), A4 with a sharp sign (#), B4 with a flat sign (b), and C5. Frase 3 consists of four notes: G4, A4, B4, and C5. There is a significant gap between the end of Frase 2 and the beginning of Frase 3, indicating a large separation between the two phrases.

, **Figura 5.19:** Frase 2 y 3 de la melodía en progreso con una distancia de separación grande.

El resultado de la transposición de las frases se muestra en la figura 5.18. Con la modificación en la segunda frase de la melodía en progreso, la coherencia que antes existía entre la frase 2 y la frase 3 se ha roto. La figura 5.19 muestra como han quedado las frases 2 y 3 después de la modificación.



, *Figura 5.20: Resultado de la segunda modificación hecha por reflection*



, *Figura 5.21: Resultado de la tercer modificación hecha por reflection*

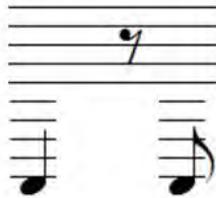
Nuevamente el sistema detecta el salto grande que hay entre las frases y debe modificarlo para mantener la coherencia. En esta ocasión la diferencia es de 8 tonos y se seleccionó aleatoriamente una transposición para corregir la inconsistencia. Esta vez la transposición es de 6 tonos. El resultado se muestra en la figura 5.20.

La coherencia que existía entre las frases 3 y 4 de la melodía en progreso se rompió cuando la frase 3 fue movida. La diferencia que el sistema encuentra que existe entre las frases es de $9\frac{1}{2}$ tonos, nuevamente un salto grande, pues hay que recordar que el usuario permite saltos de máximo 6 tonos. Para corregir esta inconsistencia se ha seleccionado aleatoriamente nuevamente una transposición y en esta ocasión la transposición es de 7 tonos. La figura 5.21 muestra el resultado después del estado *reflection*. Las modificaciones hechas en la melodía cambian mucho a la misma, de hecho aunque las frases utilizadas forman parte de las melodías previas una vez unidas y modificadas, hacen poca o casi ninguna referencia a su origen.

De esta manera termina la primera ejecución del ciclo E-R completa.



, **Figura 5.22:** *Primer frase agregada en el segundo ciclo engagement. Frase 5*



, **Figura 5.23:** *Segunda frase agregada en el segundo ciclo engagement. Frase 6*

5.2.3. Segunda ejecución de *engagement*

Los parámetros que se utilizaron para la primera ejecución no cambian. La segunda ejecución de *engagement* toma la melodía en el punto en que la dejó *reflection*. El contexto de la melodía se actualiza y el ciclo inicia.

El procedimiento es el mismo que en el ciclo anterior y los resultados se muestran a continuación.

Primer ciclo *Engagement*

Cantidad total de Átomos similares o iguales encontrados: 54

El átomo seleccionado aleatoriamente fue: Atomo430

La frase incorporada se muestra en la figura 5.22:

Segundo ciclo *Engagement*

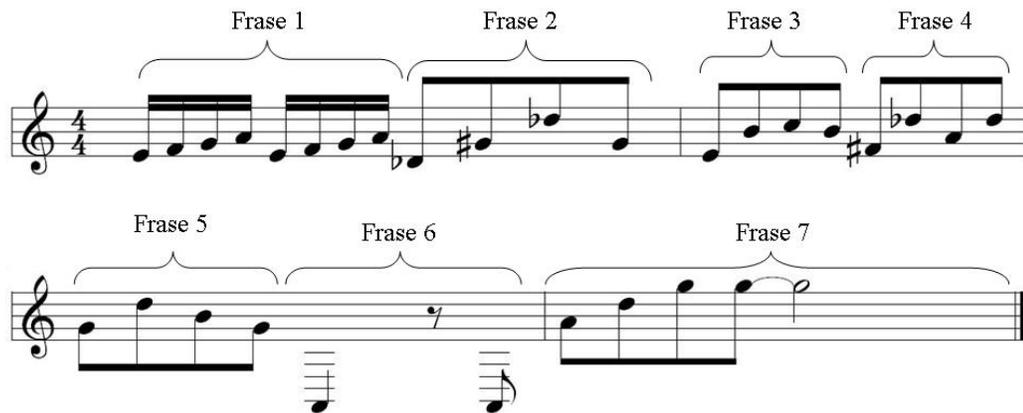
Cantidad total de Átomos similares o iguales encontrados: 72

El átomo seleccionado aleatoriamente fue: Atomo637

La frase incorporada se muestra en la figura 5.23:



, *Figura 5.24: Tercer frase agregada en el segundo ciclo engagement. Frase 7*



, *Figura 5.25: Melodía en progreso después de la segunda ejecución de engagement*

Tercer ciclo *Engagement*

Cantidad total de Átomos similares o iguales encontrados: 40

El átomo seleccionado aleatoriamente fue: Atomo327

La frase incorporada se muestra en la figura 5.24:

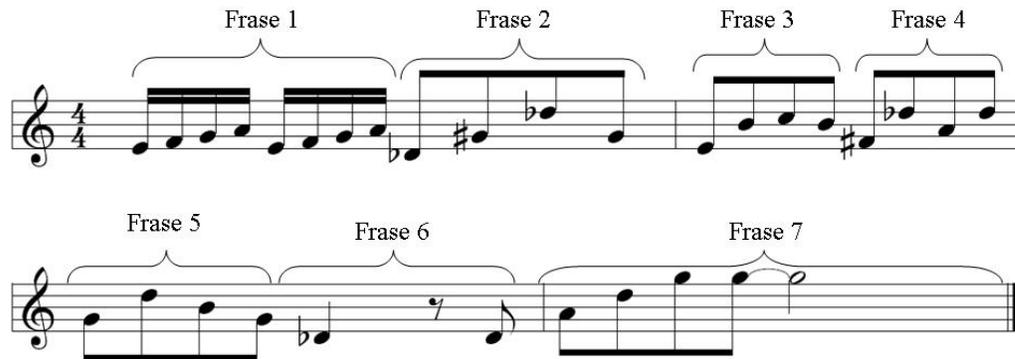
La melodía en progreso hasta el momento se muestra en la figura 5.25.

5.2.4. Segunda ejecución de *reflection*

Los parámetros que se utilizaron en la primera ejecución de *reflection* tampoco cambian. El procedimiento es el mismo y los resultados obtenidos se muestran a continuación.

Los valores de copia en la melodía encontrados por el sistema fueron:

Porcentaje de copia de melodía	15.625 %
Porcentaje de copia del ritmo	39.39393939393939 %



, *Figura 5.26: Melodía después de la segunda ejecución del ciclo E-R completo*

Con esos valores, de la misma forma que en el caso anterior, el sistema establece que la melodía generada no es una copia de las melodías previas.

Entre las frases 5 y 6 recién agregadas a la melodía en progreso existe una distancia que es mayor que la permitida por el parámetro *Máxima Diferencia entre notas*. La distancia es de 11 tonos y el procedimiento seleccionado por el sistema para resolver el problema es una transposición que será de 8 tonos para acercar la frase 6 a la 5.

Luego de mover la frase 6 la distancia con la frase 7 se reduce quedando dentro del valor permitido, de tal forma que no hay que modificar más.

El resultado de la melodía en progreso se muestra en la figura 5.26.

5.2.5. Tercera ejecución de *engagement*

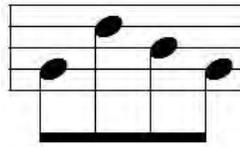
Los parámetros que se utilizaron para la primera y segunda ejecución no cambian. La tercera ejecución de *engagement* toma la melodía en el punto en que la dejó *reflection*. El contexto de la melodía se actualiza y el ciclo inicia. El procedimiento es el mismo que en el ciclo anterior y los resultados se muestran a continuación.

Primer ciclo *Engagement*

Cantidad total de Átomos similares o iguales encontrados: 48

El átomo seleccionado aleatoriamente fue: Atomo438

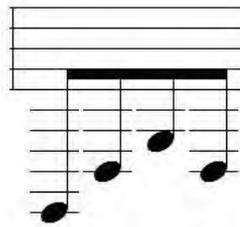
La frase incorporada se muestra en la figura 5.27:



, **Figura 5.27:** *Primer frase agregada en el tercer ciclo engagement. Frase 8*



, **Figura 5.28:** *Segunda frase agregada en el tercer ciclo engagement. Frase 9*



, **Figura 5.29:** *Tercer frase agregada en el tercer ciclo engagement. Frase 10*

Segundo ciclo *Engagement*

Cantidad total de Átomos similares o iguales encontrados: 80

El átomo seleccionado aleatoriamente fue: Atomo440

La frase incorporada se muestra en la figura 5.28:

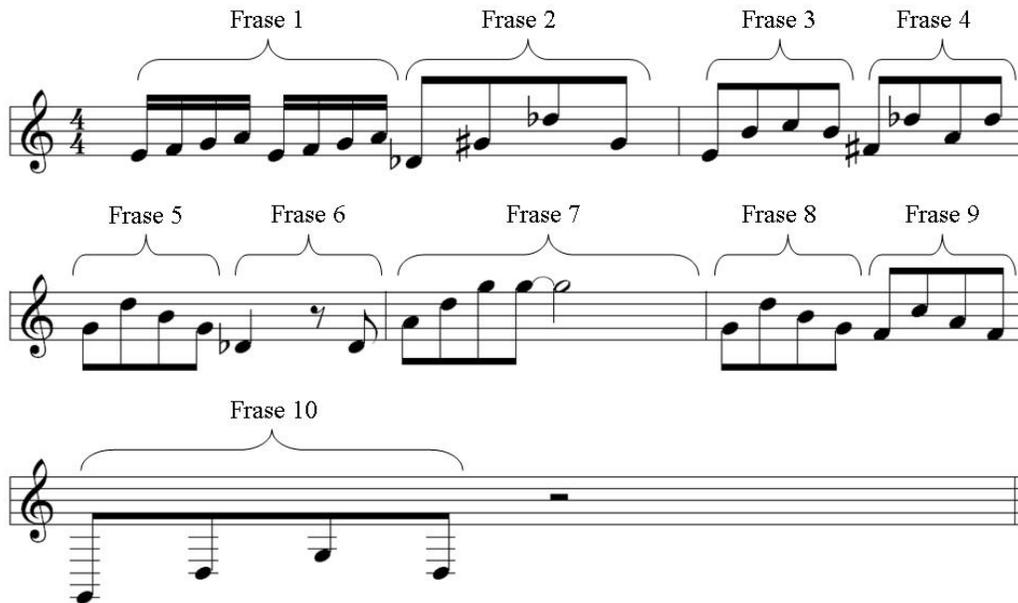
Tercer ciclo *Engagement*

Cantidad total de Átomos similares o iguales encontrados: 114

El átomo seleccionado aleatoriamente fue: Atomo373

La frase incorporada se muestra en la figura 5.29:

La melodía en progreso hasta el momento se muestra en la figura 5.30.



, *Figura 5.30: Melodía en progreso después de la tercera ejecución de engagement*

5.2.6. Tercera ejecución de *reflection*

Los parámetros que se utilizaron en la primera y segunda ejecución de *reflection* tampoco cambian. El procedimiento es el mismo y los resultados obtenidos se muestran a continuación.

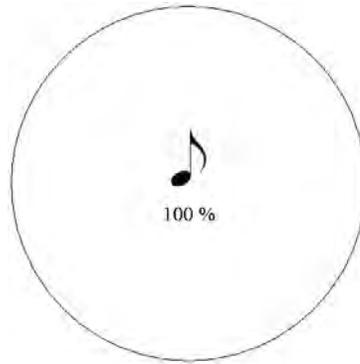
Los valores de copia en la melodía encontrados por el sistema fueron:

Porcentaje de copia de melodía	9.523809523809524 %
Porcentaje de copia del ritmo	43.18181818181818 %

Con esos valores, de la misma forma que en el caso anterior, el sistema establece que la melodía generada no es una copia de las melodías previas y no es necesario modificar parámetros del sistema.

Entre las frases 9 y 10 existe una diferencia de 11 tonos, la máxima diferencia permitida por el usuario es de 6 tonos y por lo tanto hay que modificar. El procedimiento seleccionado por el sistema para resolver el problema es la construcción de un *spline* para unir las dos frases. Este procedimiento no se había utilizado, así que se explica con detalle.

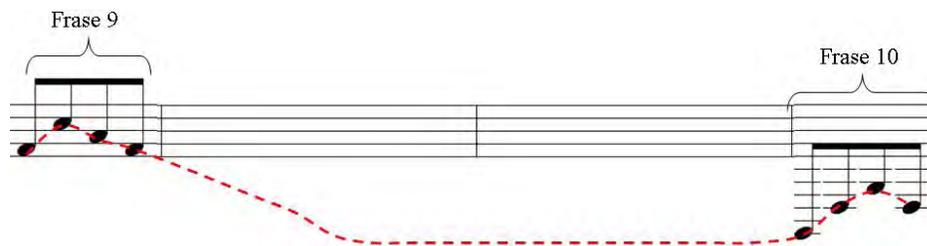
Como se mostró en la sección 4.4.2 lo primero que se debe hacer es construir la ruleta de probabilidades para seleccionar las duraciones de las notas inter-



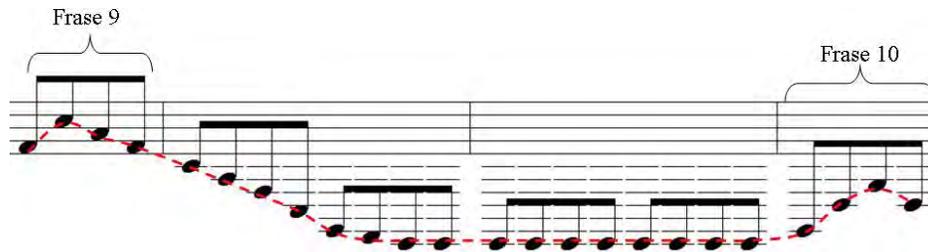
, **Figura 5.31:** Ruleta de probabilidades para seleccionar las duraciones de las notas en el spline.

poladas. Con este procedimiento también se establece el número de notas a interpolar. Para este caso las notas en la frases 9 y 10 son todas corcheas, así que la probabilidad de seleccionar una corchea para insertarla entre las dos frases es 100 %. En la figura 5.31 se muestra la ruleta.

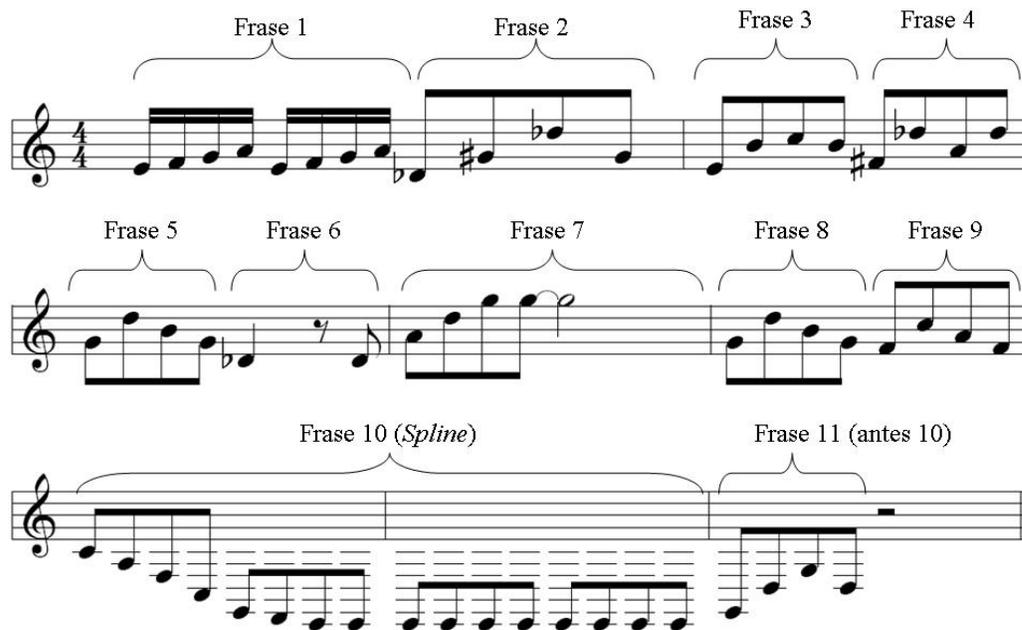
Después se genera la curva y las notas ajustan su altura (tono) al contorno de esta. En la figura 5.32 se muestra la curva generada y en la figura 5.33 se muestran la frase 9, la interpolación de notas y la frase 10. Las notas interpoladas constituyen una nueva frase en la melodía y es nombrada frase 10, la anterior frase 10 es renombrada como frase 11. En este ejemplo se puede observar que las notas en la parte baja de la curva no cambian. Las melodías generadas por el sistema tienen como fin ser ejecutadas en una guitarra, la nota más grave que se toca en este instrumento es un Mi de la 3 octava (E3) y esa es la nota en la que la curva y las notas ya no bajan más, pues de otra forma no podrían tocarse. Esta es una limitación impuesta por el autor.



, **Figura 5.32:** Curva (spline) generada entre las frases 9 y 10 para interpolar notas.



, *Figura 5.33: Frases 9 y 10 y las notas interpoladas entre ellas que se ajustan a la curva generada.*



, *Figura 5.34: Melodía resultante después de tres ciclos E-R completos.*

No existen más saltos grandes que rompan la coherencia de la melodía. De esta forma termina la tercera ejecución del ciclo *engagement-reflection*. La melodía resultante se muestra en la figura 5.34.

5.3. Limitaciones

ERMEG no siempre da resultados satisfactorios. Existen varias causas para que el sistema falle en la composición de una melodía. Por ejemplo, cuando la constante de resolución tiene valores de: 100 % para la resolución local y 100 % para la global, sucede que el sistema busca encontrar similitudes con todos los puntos de las gráficas de emoción y tensión de 100 %, es decir, gráficas iguales. Esto es poco probable que suceda, y si sucede, el sistema con mucha probabilidad copiará una melodía previa. Cuando por ejemplo, la constante de resolución local tiene un valor de 0 % y la global de 100 % el sistema quita toda la importancia a la comparación del valor local de emoción y tensión y concentra la evaluación solo los valores del vector histórico. Aunque considerar los aspectos pasados está bien, no lo es todo, pues el valor actual indica el último giro que ha tomado la melodía, y este puede ser muy importante al momento de agregar nuevas frases. Puede suceder que la constante de resolución local tenga un valor de 100 % y la global de 0 %, en este caso el usuario quita la importancia al vector histórico de tensiones y emociones y concentra la evaluación en los valores locales. De esta forma solo importa el valor actual de emoción y tensión, y todo lo que ha pasado antes en la melodía carece de importancia, cayendo en el error de otros sistemas de composición automática que no toman en cuenta el contexto de la melodía. La constante ACAS tiene problemas similares. Si los parámetros ACAS superior e inferior tienen valores de 100 % el sistema sólo considera como posibles opciones para continuar la melodía aquellos átomos con una similitud al contexto actual de la melodía de 100 % (rango entre 100 % y 100 %). Encontrar átomos 100 % similares sucede generalmente en las primeras ejecuciones de los ciclos E-R cuando la melodía en progreso es pequeña y son pocos los puntos de tensión y emoción que se comparan, pero tiende a dejar de suceder conforme crece el número de frases que la forman, con lo que rápidamente se llega a un *impass* y el programa termina. Si el programa no termina su ejecución es por que se está copiando una melodía previa, por eso le es posible encontrar átomos iguales después de varias ejecuciones de ciclos E-R. Si el caso fuera que las constantes ACAS tienen valores de 100 % la superior y 0 % la inferior, el sistema ya no toma en cuenta la similitud de los contextos, pues se están buscando similitudes entre el 100 % y el 0 % y cualquier átomo cumple con este criterio. De esta forma el sistema selecciona como posibles opciones para continuar la composición a todos los átomos de la base de conocimiento, y como la selección de la opción para continuar la melodía es

aleatoria el proceso de composición se convierte en una selección aleatoria de la siguiente frase. También puede encontrarse el caso de que las constantes tengan valores de 0 % la superior y 0 % la inferior con lo que el sistema estaría buscando los átomos que no se parezcan en absoluto al contexto actual de la melodía, lo cual no es común y además no parece razonable.

Aún con lo descrito, la manipulación de los valores de las constantes ACAS y de resolución puede generar resultados interesantes en las composiciones del sistema.

Otra situación en la que el sistema no da resultados satisfactorios es cuando las frases composicionales (FC) son muy grandes o muy pequeñas. Si son muy grandes, los fragmentos de las melodías previas que forman las FC son fácilmente identificables y no es un objetivo del sistema que se identifique la fuente de sus composiciones. Si las FC son muy pequeñas no es posible identificar de cual melodía previa provienen, pero también es difícil encontrar algún significado musical en ellas. Por ejemplo, si las frases fueran de una sola nota no expresan mucho por sí mismas y no es posible o no tiene sentido calcular velocidad, diferencia interválica o variación interválica, y aún menos calcular algún valor de emoción o tensión.

5.4. Discusión

En este capítulo hemos visto un ejemplo de como trabaja ERMEG. Se ha mostrado que el sistema es capaz de componer melodías siguiendo el modelo E-R. Se vio el modo en el que interactúan diversos parámetros del sistema para conducir la composición. También se mostró como funcionan todos los procesos de los estados *engagement* y *reflection*.

En el siguiente capítulo se hará una evaluación del sistema y de sus resultados. Esta evaluación servirá para determinar si el uso del modelo E-R en composición automática puede generar resultados notables como lo ha hecho en la generación de historias.

Capítulo 6

Evaluación

6.1. Introducción

En éste capítulo se presenta una evaluación de las metas que se habían fijado al inicio de este trabajo de tesis. Tales metas se desarrollaron como rutinas en el sistema y se explicaron en el capítulo 4. También se presentaron ejemplos de su ejecución en el capítulo 5, pero ahora se examinan con detalle para determinar que logran resultados adecuados. Además de lo anterior se presentan opiniones de usuarios del sistema sobre las salidas arrojadas por él. Estas opiniones sirven de retroalimentación en esta investigación para saber qué aspectos tienen resultados aceptables y cuáles se deben mejorar.

Para evaluar las rutinas implementadas que llevan a la consecución de una meta se realizan varios ejemplos para mostrar la forma de operar del sistema en distintas situaciones. Con los ejemplos mostrados se podrá deducir que se han cumplido las metas fijadas. Las opiniones de usuarios se recolectan mediante un cuestionario y se extrae de sus respuestas información de retroalimentación para la mejora del sistema.

Es importante analizar los aspectos anteriores para saber que el trabajo realizado cumple con los requisitos fijados y para saber hacia dónde partir en investigaciones futuras.

6.2. Evaluación de las metas del sistema

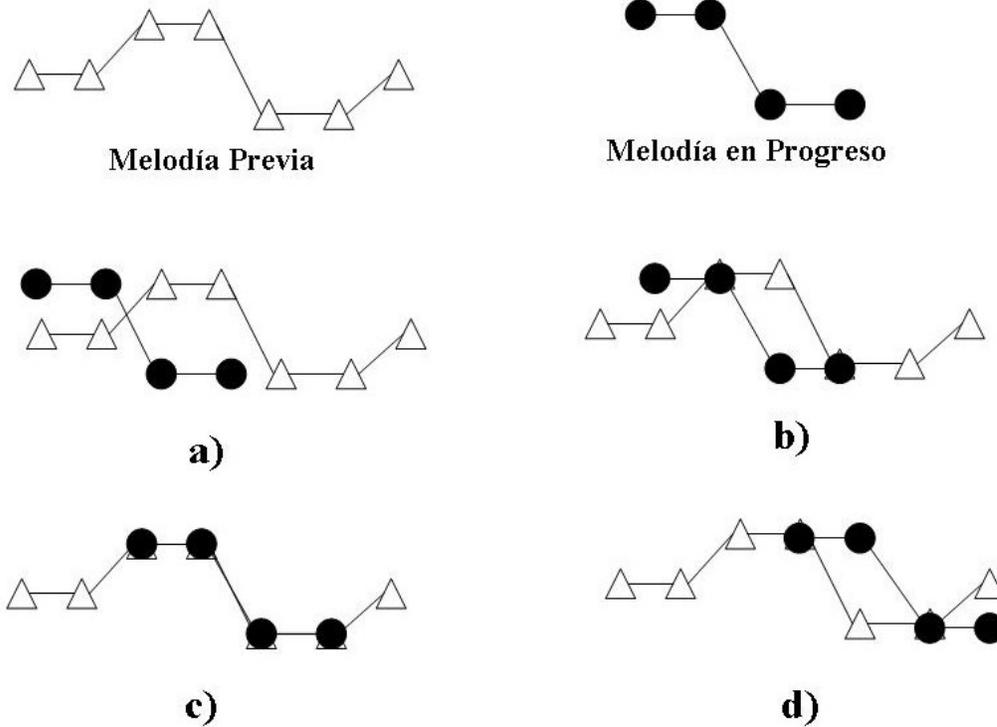
ERMEG tiene varias metas fijadas en este trabajo. En este apartado se evalúa si éstas se logran y cuál es la confiabilidad de su ejecución, es decir, si

cumplen siempre su cometido o sólo a veces. Dentro del conjunto de metas hay dos que son muy importantes en los resultados que arroja el sistema. Estas metas son percibidas por los usuarios del sistema y le dan sentido al proceso realizado por el sistema, ellas son: la novedad y la coherencia de la composición. Estas características son subjetivas, pero el sistema integra una definición por medio de la cual afirmar si cumplen con tales criterios. Existen otras metas que tiene que ver con la manera de enlazar ideas que tiene el sistema utilizando para ello una representación de emociones y luego las estructuras de átomos y contextos.

6.2.1. Evaluación de la Novedad

Como se mencionó en la sección 2.1.3 una parte muy importante de los sistemas creativos, para que sean aceptados como tales, es que los resultados deben ser novedosos. Siguiendo estas ideas, para los sistemas computacionales se ha propuesto la idea de la creatividad computacional [Pérez y Pérez & Sharples, 2004], en la que la novedad se mide con respecto a la base de conocimiento del sistema. Un sistema será creativo si es capaz de producir resultados que no están previamente en su base de conocimiento. ERMEG incorpora la capacidad de producir resultados diferentes a los de su base de conocimiento. Para lograrlo utiliza un algoritmo de comparación de las melodías en progreso contra todo el conjunto de melodías previas. En la sección 4.4.2 se explicó la forma en que trabaja este algoritmo comparando punto a punto las gráficas de melodía y ritmo. De la comparación surgen valores de semejanza en varios niveles de la melodía. En la figura 6.1 se puede ver un ejemplo de la búsqueda de similitudes. La mayor similitud es del 100 %, y es esa la que se guarda para que el sistema la compare contra el parámetro de porcentaje máximo de copia ajustado por el usuario del sistema. En la imagen 6.1 cada punto o triangulo representan un evento de la melodía, ya sea una nota en la gráfica de melodía, o una duración en la gráfica de ritmo. En este ejemplo mostrado existen muy pocos eventos pero típicamente, aun en composiciones cortas, se tienen muchos más. De tal forma que no es fácil encontrar similitudes tan altas a menos que se estén copiando fragmentos de melodías previas.

La estrategia que se sigue para evitar que el sistema copie las melodías previas es que se decrementa el valor de las constantes de resolución y ACAS para encontrar átomos novedosos que asociar. De esta forma se logra que el sistema disponga de más átomos entre los cuales buscar una opción para continuar



, **Figura 6.1:** La melodía en progreso se compara sobre todos los niveles de la melodía previa. Cuando existen similitudes entre eventos estos se traslapan. En a) no existe ningún traslape la similitud en esa sección de la melodía previa es 0 %. En b) Existen dos traslapes con lo que la similitud de la pieza en progreso y esa sección es del 50 %. En c) todos los puntos de la melodía en progreso se traslapan con esa sección de la melodía previa, la similitud es del 100 %. En d) nuevamente hay dos traslapes y la similitud es del 50 %.

la composición; uno de ellos lo llevan a seguir copiando la melodía previa, los demás no. La selección de la siguiente frase musical es un proceso aleatorio con lo que si se incrementan las opciones de frases a elegir existe una alta probabilidad de que no se elija la que continúa copiando una melodía previa. Puede suceder que el sistema seleccione la que continúa copiando, entonces el sistema encuentra que otra vez el porcentaje de copia de la melodía esta por encima del máximo permitido por el usuario y nuevamente decrementa los valores de las constantes. Lo anterior provoca que aún más átomos sean considerados para la selección de la siguiente frase, con lo que la probabilidad de seleccionar la única frase que continua la copia de una melodía previa es menor. De esta forma se trata de forzar al sistema a buscar opciones alternas para continuar la composición sin copiar las melodías previas. Es de suponerse que aunque la probabilidad de seleccionar la frase que continúa la copia se reduzca cada vez más la copia podría suceder de cualquier modo. Con lo que no se puede asegurar que el sistema siempre producirá melodías novedosas, pero se puede decir que con una alta probabilidad lo hará.

Es importante mencionar que el sistema puede retroalimentarse con las melodías que genera, pero que al final de cuentas utilizan las mismas frases composicionales (excepto la frase inicial y las modificaciones hechas por reflection: transposiciones y splines). Estas frases podrían encontrarse en el proceso de selección cuando el sistema estuviera copiando una melodía previa, con lo que serían dos los átomos que al seleccionarse continuarían la copia. El número de átomos que continuarían la copia y que podrían considerarse en una selección se incrementará como se incrementen las melodías retroalimentadas del sistema. Aún así la probabilidad de seleccionar alguna de tales frases es pequeña, más aún la probabilidad de seguirlo haciendo en selecciones posteriores es también pequeña. De tal modo que podemos decir nuevamente que con alta probabilidad el sistema producirá melodías novedosas.

Una última consideración que se debe hacer al respecto es que el sistema si no enriquece constantemente su base de conocimiento no podría generar melodías nuevas para siempre. Esto se debe a que, por ejemplo, en la ejecución paso a paso vista en el capítulo 5 se tenían 699 átomos, cuando se lograrán combinar todos ellos de todas las formas posibles en nuevas melodías, sería imposible generar otra más que fuera distinta a cualquiera de ellas. El número de posibles combinaciones es muy grande y alcanzarlo, incluso con esta base de conocimiento pequeña, resulta muy difícil. Además si la base de conocimiento se enriquece resulta más difícil que esto suceda.

En conclusión, si la base de conocimiento no se enriquece con nuevas melodías

(que no genere el sistema) ERMEG no puede componer para siempre algo novedoso. Pero aún sin actualizar la base de conocimiento es posible generar un número muy grande de melodías previas con una alta probabilidad.

6.2.2. Evaluación de la Coherencia

Una meta muy importante en el sistema es la de construir una melodía que sea coherente. La coherencia tuvo que ser definida para el sistema y se tiene que una melodía lo es cuando no existen saltos grandes entre las notas que separan las frases que han sido puestas juntas. El salto se mide con respecto a la nota final de una frase y la inicial de la siguiente. El salto será grande si supera el máximo número de tonos de separación definido por el usuario del sistema.

La estrategia que sigue el sistema para resolver problemas de incoherencia en las melodías en progreso se describió en la sección 4.4.2 y se ejemplificó en la sección 5. El procedimiento es un algoritmo determinista que ejecuta las modificaciones a la melodía en progreso siguiendo siempre una secuencia de pasos bien definida por lo que siempre se puede saber que es lo que ocurrirá. Las modificaciones son realizadas por dos procedimientos que son seleccionados aleatoriamente. Uno de ellos es aplicar una transposición proporcional a la distancia del salto que separa las frases. El otro es construir un *spline* natural e interpolar notas en dos compases que unen las dos frases sin modificarlas.

En ambos casos pueden ocurrir problemas cuando las distancias que separan las frases son enormes. En una transposición, cuando algo así sucede, el sistema ejecuta un movimiento de la segunda frase una cantidad proporcional al salto que separa las frases. La proporción de ese movimiento es definida por el usuario. Si la separación es muy grande y la proporción del movimiento es pequeña probablemente la diferencia entre ambas seguirá siendo superior al parámetro de máxima distancia fijado por el usuario, con lo que la melodía continuará siendo no coherente. Tal problema requerirá varias ejecuciones de transposición para solucionar la coherencia. En la construcción del *spline* si las distancias que separan las frases son muy grandes las notas colocadas sobre la curva tendrán separaciones también grandes para lograr unir las dos frases sin modificarlas. Esto no representaría que la melodía no es coherente y aún así podría escucharse un poco mal o ser un poco difícil de tocar. En la figura 6.2 se muestran ejemplos de lo anterior. En la figura, en *a)* se representa las dos frases originales que fueron puestas juntas por el siste-

ma. La distancia que las separa es de 48 tonos y siguiendo los parámetros indicados en el capítulo 5, la distancia máxima permitida (6 tonos) ha sido superada ampliamente por lo que la coherencia se ha roto. En *b*) la estrategia fue recuperar la coherencia con una transposición y nuevamente utilizando el parámetro de la proporción de salto usado en los ejemplos del capítulo 5 que es $2/3$, la magnitud del salto es de 32 tonos (un salto muy grande), pero aún así la nueva distancia que separa a las frases es de 16 tonos y sigue rompiendo la coherencia. Una sucesiva ejecución de una transposición reduciría la distancia a 5.5 tonos, con lo que se logra la coherencia de la melodía. En *c*), con notas interpoladas sobre la curva de un *spline* natural, se trata de conseguir la coherencia de la melodía. En este caso de inmediato se recupera la coherencia, no hacen falta más ejecuciones, pero los “pasos” entre notas interpoladas son un poco grandes lo que complica un poco la ejecución y puede comprometer la calidad de la melodía.

Aun requiriendo de más de una ejecución de las rutinas de *reflection* (lo que típicamente sucede), para lograr la coherencia en la melodía el sistema siempre es capaz de lograrla.

De cualquier forma es poco probable que sucedan saltos entre frases así de grandes pues la selección de las frases que se colocan juntas se basa en los niveles de tensión y emoción presentes en ellas. Estos a su vez se basan en parámetros como la octava en la que se tocan las frases, con lo que la selección esta correlacionada con la octava de las frases, así que no es normal encontrar separaciones de frases tan grandes.

6.2.3. Otras Metas del Sistema

Las anteriores metas son importantes, y se pueden apreciar en las melodías producidas por el sistema, pero existen otras metas definidas en el capítulo 1 que son importantes durante el proceso de composición que lleva a cabo ERMEG. Estas incluyen la generación de una representación de emociones y tensiones, construcción de algoritmos y rutinas para lograr ligar ideas musicales relacionadas por sus representaciones de emoción y tensión y emplear el conjunto de todas las emociones y tensiones presentes en las melodías para ligar las ideas nuevas, es decir, emplear el contexto de las melodías, que es precisamente una de las mayores fallas de sistemas de composición musical que emplean un sistema de cómputo. Se han alcanzado tales metas y el detalle de cada una de ellas se ha mencionado largamente durante el capítulo 4. En el capítulo 5 se mostró la forma en que trabajan en conjunto y los re-

The figure consists of three parts, a), b), and c), each showing musical notation on a 4/4 staff. Part a) shows two phrases, 'Frase 1' and 'Frase 2', on separate staves. Part b) shows 'Frase 1' and a modified 'Frase 2' on the same staff, with the second phrase transposed. Part c) shows 'Frase 1' and 'Frase 2' on the same staff, with a 'Spline' curve connecting them to create a smooth transition.

, **Figura 6.2:** En a) se muestran las dos frases originales puestas juntas por el sistema. En b) se ha aplicado una transposición a la segunda frase para acercarla a la primera. La distancia que las separa aún rompe la coherencia. En c) se ha aplicado un spline para interpolar notas y la coherencia se logra.

sultados a los que conducen. Estas metas están íntimamente ligadas y forman parte del núcleo del ciclo engagement.

6.3. Retroalimentación

No es extraño encontrar en la literatura que se solicite a un grupo de sujetos dar su opinión sobre la salida producida por un sistema interactivo, particularmente en aquellos que modelan temas tan difíciles de evaluar como la creatividad. Por supuesto, los resultados no se pueden considerar más que una simple opinión (no hay forma de realizar algún tipo de inferencia estadística a partir de los mismos). Y mencionar que en la literatura se puede observar que es una práctica aceptada por la comunidad. Siguiendo a [Pérez y Pérez & Sharples, 2004], se ha procedido a hacer un cuestionario para conocer la opinión de un grupo de personas.

6.3.1. Opiniones

Se recolectaron opiniones sobre los resultados arrojados por ERMEG y para tal recolección se elaboró un cuestionario. De esta forma se han evaluado los resultados en sistemas creativos, por ejemplo MEXICA [Pérez y Pérez & Sharples, 2001] y MINTREL [Turner, 1994]. La entrevista completa se muestra en el apéndice A.

El siguiente no es un análisis estadístico formal, sino una retroalimentación con opiniones para averiguar si se avanzó por caminos adecuados en esta investigación. También, tomando como inspiración la prueba de Turing, se presentan en este ejercicio melodías compuestas por humanos y por el sistema a los participantes para juzgar los resultados del sistema.

6.3.2. Descripción del Cuestionario

La entrevista se presentó como parte de una investigación en generación de melodías musicales. Se recolectó información personal de los participantes como: edad, sexo, nivel de estudios u ocupación y si tenían instrucción musical, en caso de tenerla se pedía especificar. Cuatro melodías fueron incluidas en la entrevista: dos producidas por ERMEG, y dos compuestas por Metallica¹.

¹Grupo musical profesional de género rock

Todas las melodías se ejecutaron utilizando el formato MIDI². Las melodías de Metallica³, pequeños fragmentos de composiciones completas a las que se quito toda la instrumentación y armonía dejando sólo la línea melódica para que la comparación con ERMEG fuera más equitativa, fueron extraídos de solos de guitarra y también forman parte de la base de conocimiento del sistema. El hecho de que algunas composiciones habían sido generadas por un programa no fue mencionado para evitar prejuicios en las respuestas. Las melodías fueron seleccionadas por el autor como ejemplos representativos de las composiciones que el sistema produce. Las partituras de las cuatro melodías empleadas en este ejercicio se muestran en el apéndice B

El modo de aplicación de la entrevista fue como sigue:

- Se entregaron las instrucciones y un juego de 4 cuestionarios (uno por melodía) a cada participante.
- Se indicó que leyeran las instrucciones y luego verbalmente se aclaró cualquier duda sobre el mecanismo de la entrevista.
- Se mostraron, de forma continua, las cuatro melodías en un orden aleatorio. Este paso fue para que los participantes se familiarizaran con el tipo de melodías que iban a analizar.
- Se generó un nuevo orden aleatorio y con éste se presentaron las melodías una a una dejando unos minutos entre ellas para responder los cuestionarios.
- Al final se pidió a los participantes que ordenaran las melodías escuchadas según sus preferencias de la mejor a la peor y lo anotaran en la última parte de la entrevista.

El propósito de este ejercicio era conocer las impresiones que causaban las composiciones hechas por el sistema. Los aspectos que se deseaban analizar eran: la calidad de las composiciones, (i.e. buenas o malas), la calidad del compositor y la novedad de las melodías.

²Esto implica que el sonido de las melodías fue sintetizado producido por una computadora. Se empleo un timbre de piano porque suena más natural que el de guitarra en el formato MIDI

³Fragmentos de solos de guitarra de las canciones *To live is to die* y *Fade to black*

6.3.3. Descripción de los Participantes

La población que se consideró para este estudio son jóvenes estudiantes de licenciatura de la UAM Cuajimalpa. Se escogieron jóvenes universitarios porque ellos están familiarizados con la música que se empleó para las melodías previas. La muestra empleada corresponde a alumnos de un salón de clases. No había requerimientos particulares para los participantes como: nivel de educación musical o en computación. Los participantes de la entrevista fueron 17. La edad media de los participantes fue 19.8 años, con un mínimo de 18 y un máximo de 22 años, de los cuales fueron 6 hombres y 11 mujeres. Cuatro de los participantes declararon tener conocimientos musicales, las especificaciones dadas por ellos fueron: autodidacta, básicos (primer año), estudios de canto y 8 años en Yamaha con solfeo y composición.

6.3.4. Resultados

Aunque las melodías fueron presentadas a los participantes en orden aleatorio, ahora se presentan las respuestas obtenidas para cada una de las cuatro melodías en orden.

Las preguntas 1, 2 y 3 se analizan cuantitativamente.

1. ¿La melodía que acaba de escuchar le gustó?
 - a) nada
 - b) poco
 - c) más o menos
 - d) mucho
 - e) muchísimo

2. ¿La melodía escuchada fue compuesta por?
 - a) una persona que no sabe de composiciones musicales
 - b) una persona con pocos conocimientos y experiencia como compositor
 - c) un compositor amateur
 - d) un compositor profesional principiante
 - e) un compositor profesional con experiencia



, *Figura 6.3: Respuestas para la pregunta “¿La melodía que acaba de escuchar le gustó?” para la composición 1 del sistema.*

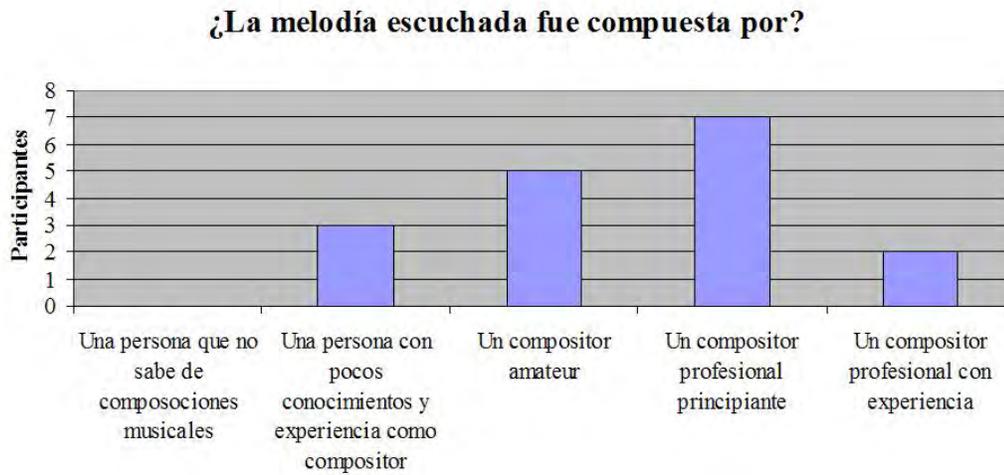
3. La composición que acaba de escuchar le parece:
 - a) una copia de otras que ha escuchado
 - b) muy similar a otras que ha escuchado
 - c) diferente a otras que ha escuchado
 - d) original
 - e) muy original

Las preguntas restantes de la encuesta se evaluarán cualitativamente. En ellas los participantes expresarán información adicional sobre las melodías escuchadas.

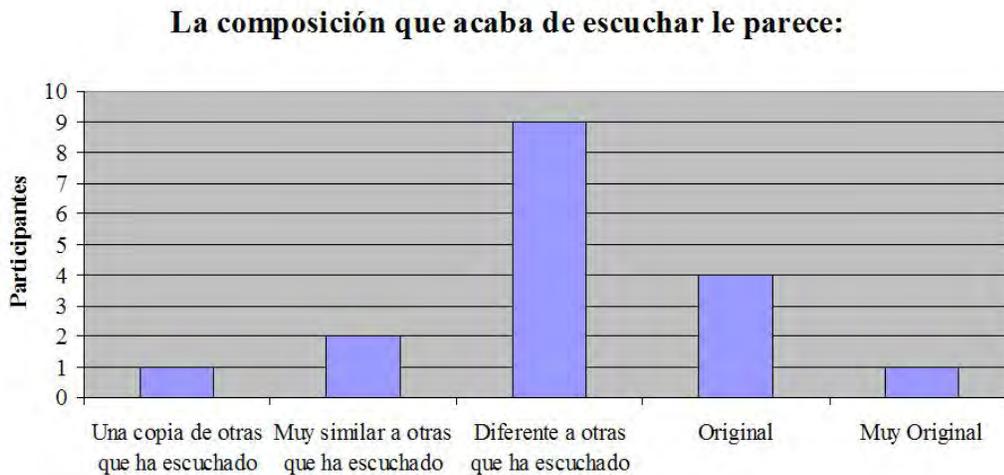
Primer melodía del sistema

La primer melodía de ERMEG obtuvo las gráficas que se muestran en las figuras 6.3, 6.4 y 6.5 para la primera, segunda y tercer preguntas respectivamente.

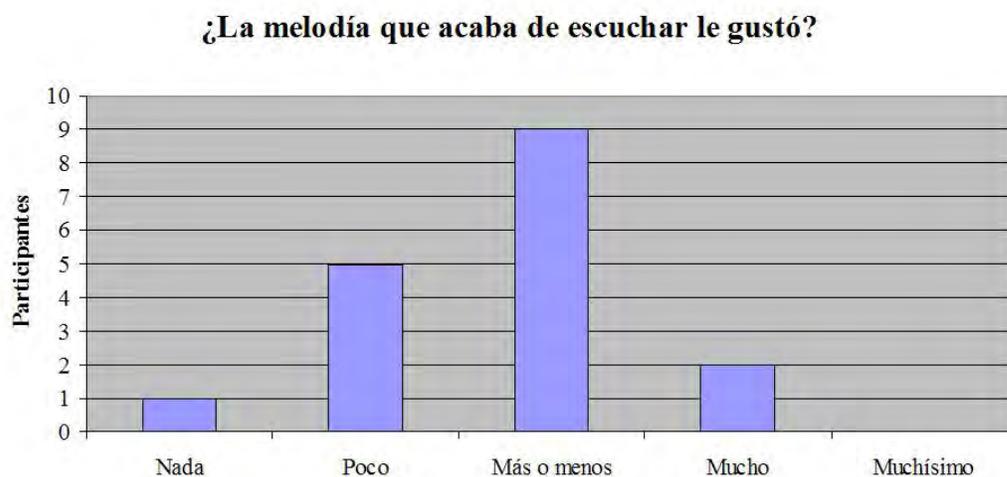
La gráfica 6.3 muestra que por una diferencia muy alta los participantes consideran que la melodía ha gustado “más o menos”. La gráfica 6.4 indica que las respuestas se concentran alrededor de tres opiniones: un compositor



, *Figura 6.4:* Respuestas para la pregunta “¿La melodía escuchada fue compuesta por?” para la composición 1 del sistema.



, *Figura 6.5:* Respuestas para la pregunta “La composición que acaba de escuchar le parece:” para la composición 1 del sistema.



, *Figura 6.6: Respuestas para la pregunta “¿La melodía que acaba de escuchar le gustó?” para la composición 2 del sistema.*

amateur, un compositor profesional principiante y un compositor profesional con experiencia con predominancia de la segunda. La gráfica 6.5 indica que la opinión general es que la melodía es: diferente a otras que han escuchado los participantes.

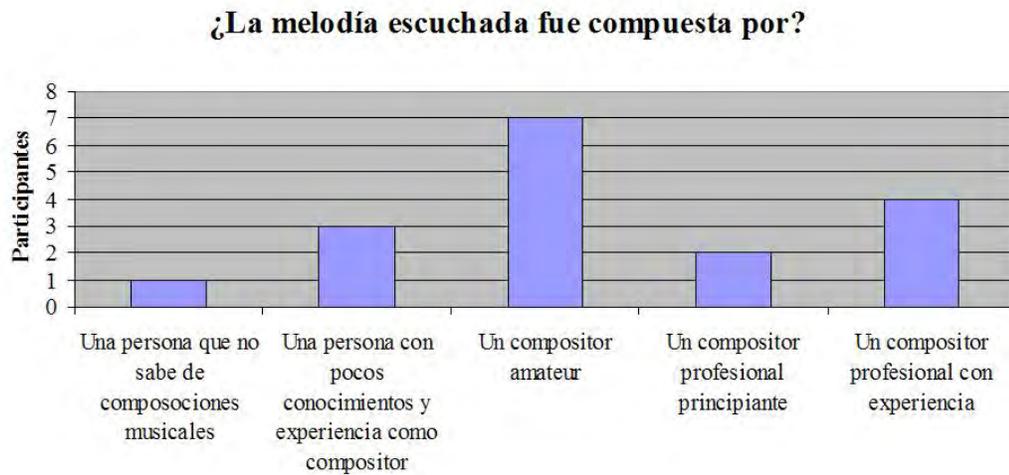
Segunda melodía del sistema

La segunda melodía de ERMEG obtuvo las gráficas que se muestran en las figuras 6.6, 6.7 y 6.8 para la primera, segunda y tercer preguntas respectivamente.

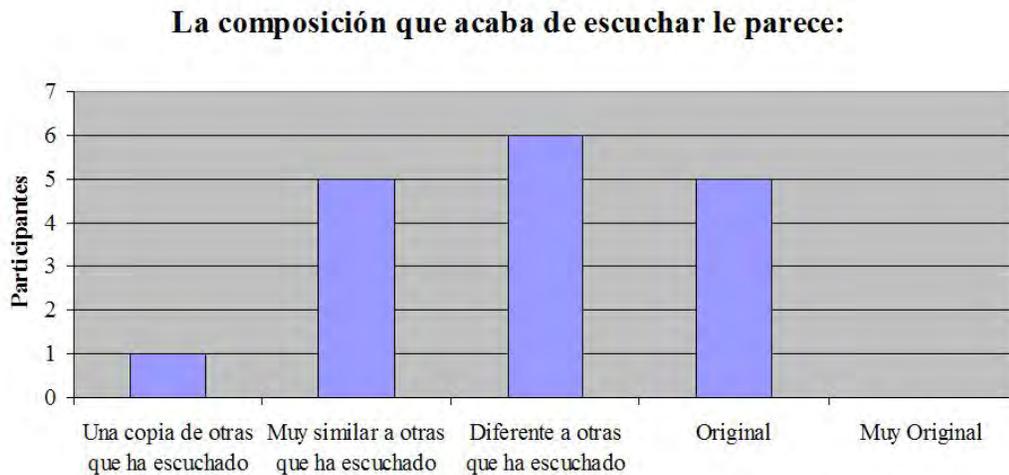
La gráfica 6.6 indica que por una diferencia alta los participantes consideran que la melodía ha gustado “más o menos”. La gráfica 6.7 muestra que la opinión mayoritaria es que la melodía fue compuesta por un músico amateur. La gráfica 6.8 indica que la opinión general es que la melodía es: diferente a otras que han escuchado los participantes.

Primer melodía de Metallica

La primer melodía de Metallica obtuvo las gráficas que se muestran en las figuras 6.9, 6.10 y 6.11 para la primera, segunda y tercer preguntas respectivamente.

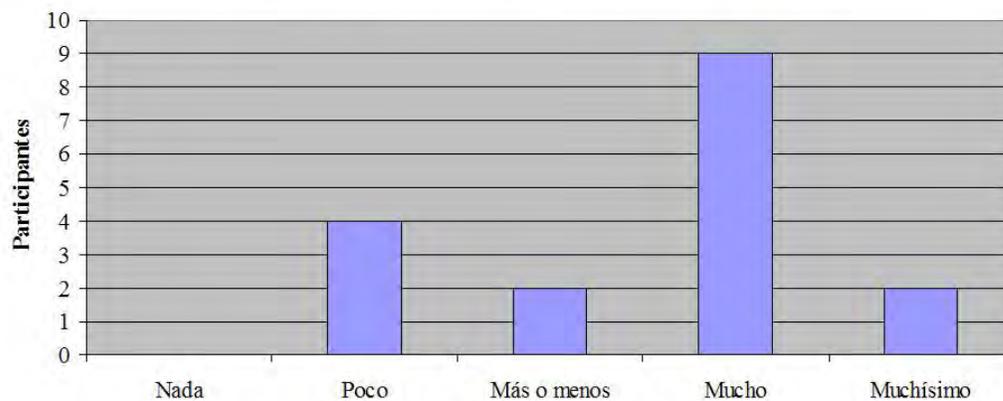


, **Figura 6.7:** Respuestas para la pregunta “¿La melodía escuchada fue compuesta por?” para la composición 2 del sistema.



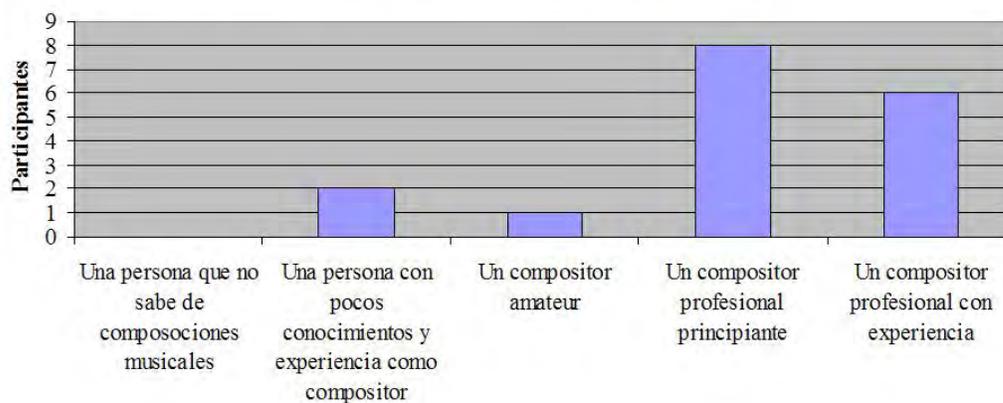
, **Figura 6.8:** Respuestas para la pregunta “La composición que acaba de escuchar le parece:” para la composición 2 del sistema.

¿La melodía que acaba de escuchar le gustó?



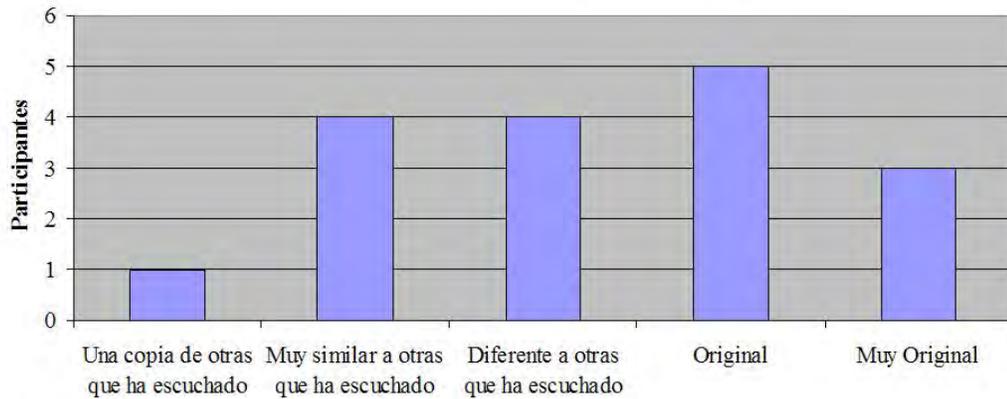
, **Figura 6.9:** Respuestas para la pregunta “¿La melodía que acaba de escuchar le gustó?” para la composición 1 de Metallica.

¿La melodía escuchada fue compuesta por?



, **Figura 6.10:** Respuestas para la pregunta “¿La melodía escuchada fue compuesta por?” para la composición 1 de Metallica.

La composición que acaba de escuchar le parece:



, **Figura 6.11:** Respuestas para la pregunta “La composición que acaba de escuchar le parece:” para la composición 1 de Metallica.

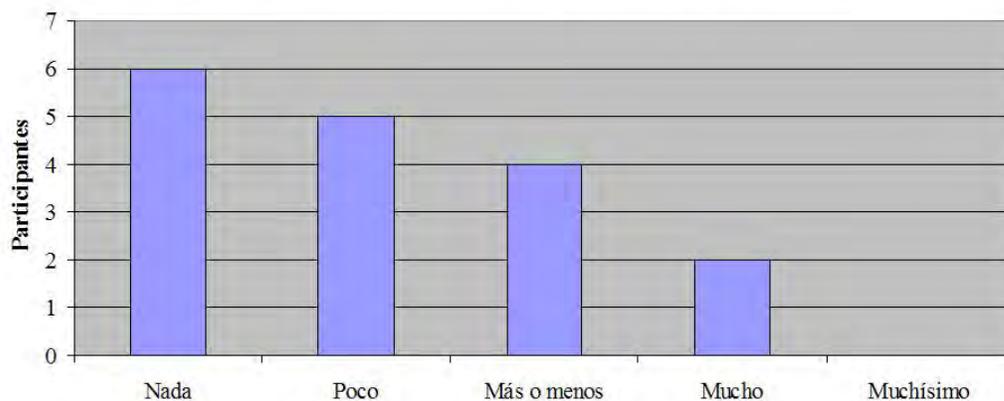
La gráfica 6.9 indica que por una amplia diferencia los participantes consideran que la melodía ha gustado “mucho”. La gráfica 6.10 muestra que la opinión mayoritaria es que la melodía fue compuesta por un compositor profesional principiante con aportación notable de la respuesta: un compositor profesional con experiencia. La gráfica 6.11 indica que la opinión mayoritaria es que la melodía es: original.

Segunda melodía de Metallica

La segunda melodía de Metallica obtuvo las gráficas que se muestran en las figuras 6.12, 6.13 y 6.14 para la primera, segunda y tercer preguntas respectivamente.

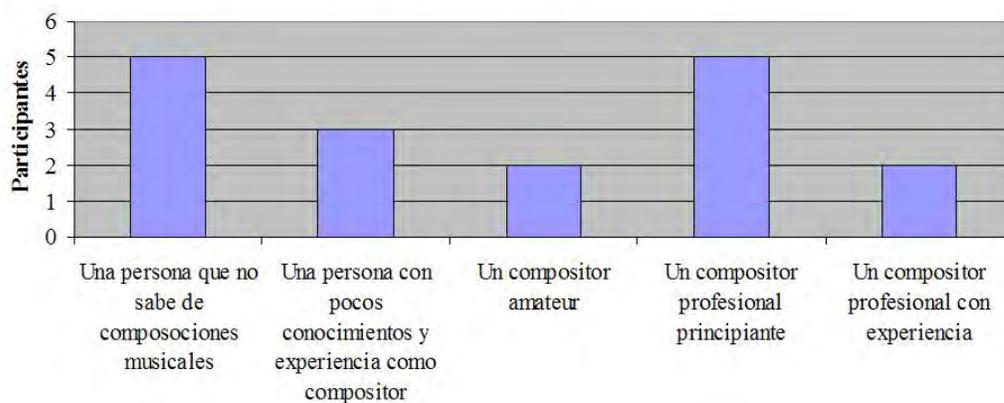
La gráfica 6.12 indica que las mayorías y la tendencia, de respuestas se encuentran en: “nada”, “poco” y “más o menos”. La gráfica 6.13 muestra una gran diversidad de opiniones que no generan una tendencia clara. Las opciones con mayoría de respuestas son: “un compositor profesional principiante” y “una persona que no sabe de composiciones musicales.” La gráfica 6.14 indica que las opiniones mayoritarias son que la melodía es: muy similar a otras que han escuchado los participantes y diferente a otras que han escuchado.

¿La melodía que acaba de escuchar le gustó?



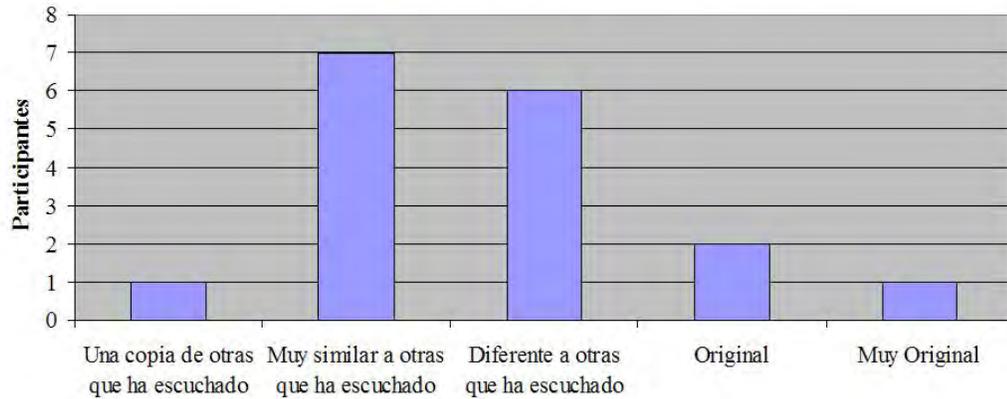
, **Figura 6.12:** Respuestas para la pregunta “¿La melodía que acaba de escuchar le gustó?” para la composición 2 de Metallica.

¿La melodía escuchada fue compuesta por?



, **Figura 6.13:** Respuestas para la pregunta “¿La melodía escuchada fue compuesta por?” para la composición 2 de Metallica.

La composición que acaba de escuchar le parece:



, *Figura 6.14: Respuestas para la pregunta “La composición que acaba de escuchar le parece:” para la composición 2 de Metallica.*

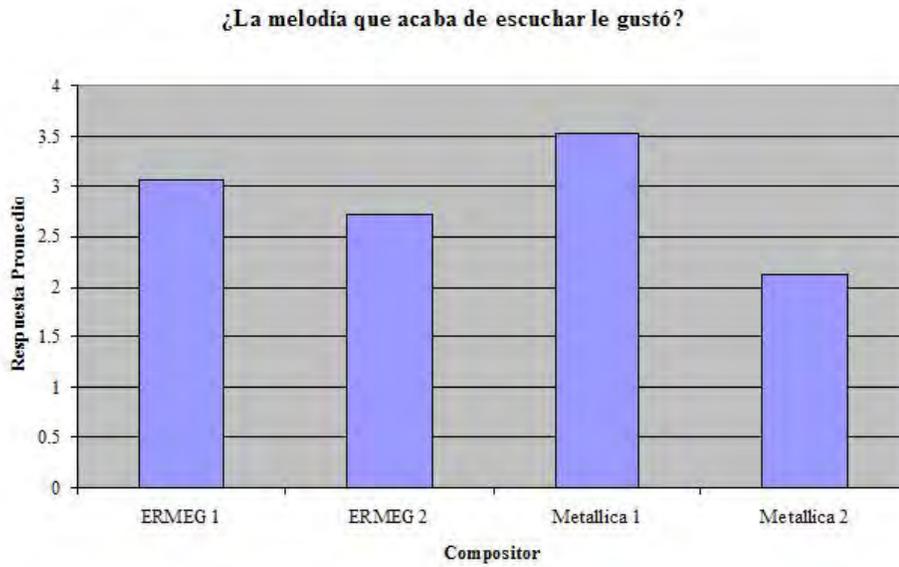
Gráficas Comparativas

En las figuras 6.15, 6.16 y 6.17 se muestran las gráficas comparativas que se obtienen para las cuatro composiciones y las tres preguntas.

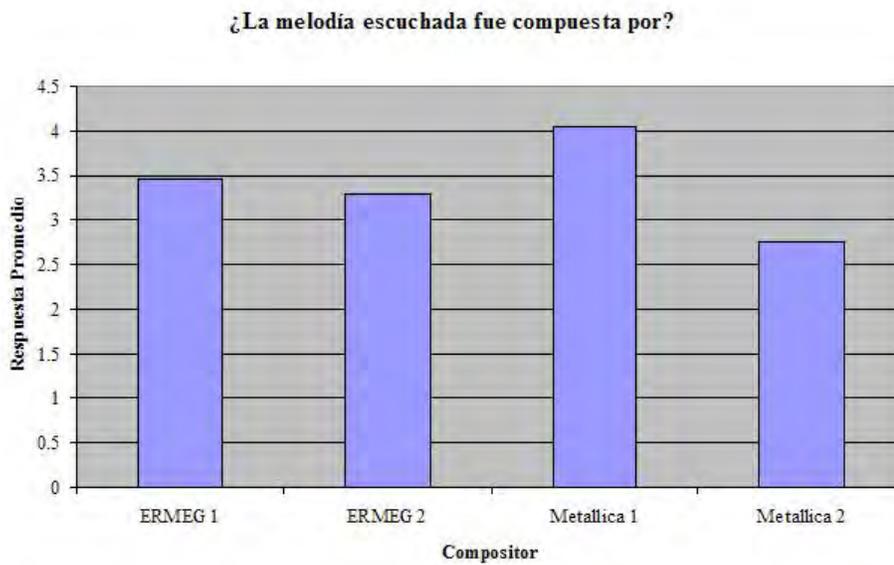
Última Parte de la Entrevista

En la última parte de la entrevista se pidió a los participantes que ordenaran las cuatro melodías escuchadas según sus preferencias. Del orden sugerido de los participantes se han obtenido los valores de mayor frecuencia de repetición en cada posición. Por ejemplo, para la primer posición (la mejor melodía), se obtuvieron múltiples respuestas, de ellas se buscó la que más se repetía y se ha colocado en la posición promedio del orden de preferencias. El mismo procedimiento se siguió con las demás melodías. El orden de preferencias promedio quedó como se muestra en la tabla 6.1.

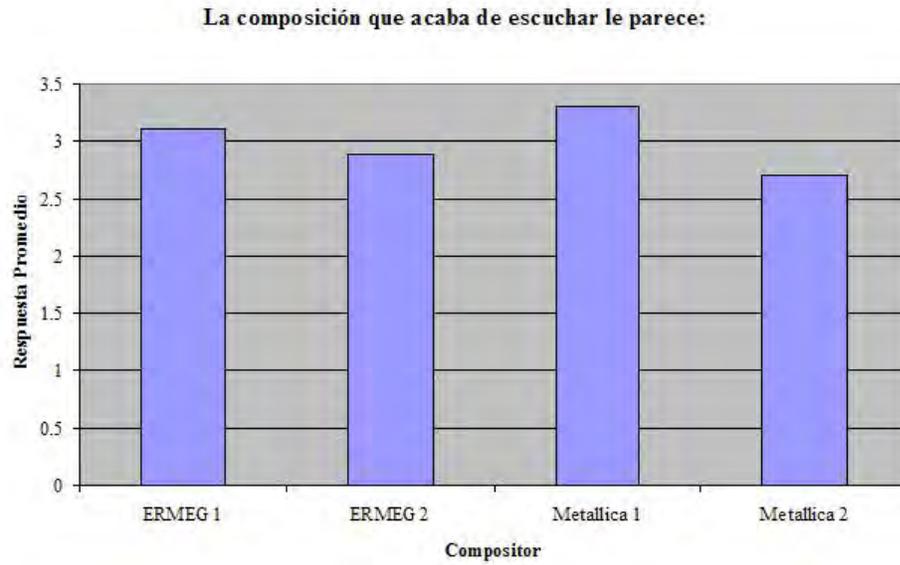
El resultado de la tabla 6.1 confirma los resultados obtenidos de las respuestas del cuestionario. En ellos se manifestó que la mejor melodía era la composición Metallica 1, y la peor fue Metallica 2. Un resultado importante de este experimento es que las melodías generadas por el sistema obtuvieron el segundo y tercer lugar. Este resultado muestra una comparación entre composiciones hechas por humanos y composiciones generadas por un programa



, *Figura 6.15: Gráfica comparativa para la pregunta 1 de las cuatro composiciones*



, *Figura 6.16: Gráfica comparativa para la pregunta 2 de las cuatro composiciones*



, *Figura 6.17: Gráfica comparativa para la pregunta 3 de las cuatro composiciones*

Posición	Melodía
#1 (la mejor)	Metallica 1
#2	ERMEG 1
#3	ERMEG 2
#4 (la peor)	Metallica 2

, *Tabla 6.1: Orden de preferencias promedio*

de computadora. Lo que sugiere este resultado no es que ERMEG es mejor compositor que Metálica, sino que para algunos de los participantes en éste ejercicio, las composiciones hechas por ERMEG tienen un nivel adecuado. En las respuestas abiertas los participantes expusieron sus impresiones personales acerca de la melodías. En lo que respecta a las melodías del sistema, los comentarios en la mayoría de los casos fueron de aceptación. Hubo comentarios como: “me parece interesante”, “es entretenida”, “tiene buen ritmo”, “interesante porque no sigue un patrón o melodía predecible”, “me pareció agradable al oído, hasta los pequeños destiempos adecuados”, etc. Algunos comentarios negativos fueron: “un poco aburrida”, “muy predecible”, “repetitiva y simple”. Con respecto a las melodías generadas por humanos los comentarios fueron: “tiene buen ritmo”, “maneja bien los silencios y el cambio de ritmo es interesante”, “en un pequeño espacio, muy breve, me recordó a un videojuego”, “está más o menos buena”, “el ritmo es bueno”, “es un poco repetitiva y sentí que a veces iba a destiempo”, “aburrida”, “interesante porque se adecua al ritmo, pero a veces algunas notas rompen con el acorde”, “me pareció una composición interesante aunque un poco repetitiva”, etc. Los comentarios hechos por los participantes confirman los resultados obtenidos. Las personas juzgan de diferentes formas a las melodías, pero en general los comentarios no son de desagrado. Los comentarios más negativos van dirigidos hacia las melodías que obtuvieron calificaciones bajas. En general las respuestas dejan ubicadas a las cuatro melodías en una posición de agrado por parte de los entrevistados.

6.4. Discusión

En este capítulo se mostró la evaluación de las metas del sistema. Tales metas fueron mencionadas en el capítulo 1 de este trabajo y desarrolladas a lo largo del mismo. Se explico el funcionamiento de las rutinas para lograr tales metas y se mostraron ejemplos. En este capítulo se hace una revisión más profunda para determinar su efectividad. También se analizaron los resultados de dos composiciones del sistema y se compararon con dos melodías compuesta por humanos por medio de una entrevista. En ella no se mencionó que dos de las composiciones habían sido generadas por un programa para no predisponer las respuestas de los participantes. Los resultados obtenidos nos llevan a pensar que las melodías compuestas por ERMEG tienen un nivel de aceptación alto entre los participantes en la entrevista.

Capítulo 7

Conclusiones

Las ciencias computacionales se han acercado a muchas disciplinas, en algunas de una forma muy natural, por la necesidad de mejorar los procedimientos y los instrumentos, por ejemplo la medicina. En otros campos la computadora ha ido ganando espacio rápidamente por ejemplo en las profesiones económico administrativas y en la industria ya que permite una mayor eficiencia en las labores que se realizan. Para el caso de las ciencias sociales también se va ganado terreno al utilizar la computadora como una herramienta indispensable, y de la misma forma lo está haciendo en las artes.

El tema de esta tesis involucra a las computadoras y a la utilidad que éstas tienen para la creación de música. Las máquinas cada vez hacen más fácil la concepción y/o arreglo de la música. A lo largo de este trabajo se ha discutido un modelo computacional para la generación automática de música. El modelo generado es una aproximación inicial y ha mostrado tener resultados promisorios. Para mejorar los resultados del sistema es necesario una investigación más profunda en el tema de música y emociones y las características que nos hacen percibirla de alguna forma, o con alguna emoción. Otra parte importante para mejorar el modelo es la inclusión de teorías musicales más completas que ayuden al sistema en el proceso de composición que ya tiene. ERMEG es el sistema producto de este trabajo de investigación y está basado en el modelo de creatividad computacional Engagement-Reflection. El sistema se compone de dos fases principales que son: la generación de las estructuras de conocimiento y un ciclo E-R para generar una melodía nueva. Era meta de esta investigación diseñar e implementar los algoritmos necesarios en el sistema para poder asociar ideas y se ha conseguido. También se desarrolló una representación de emociones y tensiones para guiar el proceso

de composición. Se desarrollaron rutinas para que el sistema sea capaz de evaluar la novedad y la coherencia de sus salidas. Estas dos características son muy importantes, pues la intención es que el sistema genere salidas nuevas (al menos con respecto a su base de conocimiento) y coherentes, lo que se traduce en melodías agradables a las personas. Una meta muy importante del sistema era contemplar el material producido para que influyera en los eventos nuevos agregados a la melodía. El sistema toma en cuenta el contexto, el cual está formado de niveles de emoción y tensión calculados para continuar la composición. La mayoría de los trabajos mencionados en el capítulo 2 no toman en cuenta el material que ya se ha producido durante la composición para generar las siguientes notas. En otros sistemas esta problemática se ha tratado de solucionar empleando cadenas de Markov. Una cadena es una serie de eventos, en la cual la probabilidad de que ocurra un evento depende de los inmediatos anteriores. Mientras más eventos anteriores se contemplen, el tamaño de la cadena crece mucho haciéndola rápidamente inmanejable. Utilizando este método y teniendo como eventos anteriores las notas que han pasado, el siguiente evento es una nota o notas con un valor de probabilidad estadísticamente tomado de otras composiciones. Otros lo han tratado de solucionar con redes neuronales pero el contexto que se considera es limitado.

El sistema, además de la ventaja de considerar el contexto de la melodía en progreso para continuar la composición, también tiene como ventaja frente a sistemas como los vistos en el capítulo 2 que no está basado en reglas musicales. La incorporación de algunas reglas podría impulsarlo de forma importante. También es una ventaja la representación de música que se emplea que ha permitido que el sistema pueda recibir la música en la base de conocimiento en cualquier tonalidad y con todas las notas de la escala musical y escribirla de la misma forma, cosa que no puede hacer la mayoría de los sistemas de composición automática.

Dentro de las limitaciones que tiene están el que no es posible agregar polifonía en una voz, ni polifonía en otras voces (más instrumentos), por lo tanto no hay armonías ni instrumentación en las composiciones del sistema. Es necesario remediar estas limitaciones. Esto se logra con el trabajo futuro de esta investigación que incluye:

1. La inclusión de armonías en un sólo instrumento (una voz) para acrecentar las posibilidades de representación de emociones y tensiones que tiene el sistema.

2. Una investigación más profunda y la inclusión de más parámetros para el cálculo de emociones y tensiones.
3. Incluir la instrumentación de las piezas generadas por el sistema para resaltar los resultados.
4. Generación de piezas más complejas que sólo líneas melódicas.

El sistema puede tener muchas aplicaciones y su contribución puede ser importante. Se puede usar para varias disciplinas pues la música da realce a varias actividades como: cine, teatro, radio, televisión, videojuegos, etc. Imaginemos un videojuego que queremos que capture mucho al usuario, si siempre presenta la misma música en cualquier situación puede volverse aburrido, si la música es generada dinámicamente de acuerdo a lo que sucede en el juego el usuario quedará más inmerso en él, y además la música no será nunca la misma.

Se mejoró el modelo E-R como resultado de la inclusión de nuevos parámetros que cambian dinámicamente y la inclusión de otros que eran necesarios para el dominio específico de la música

En la introducción de este trabajo se hicieron algunas preguntas de investigación que ahora se contestan.

- ¿Es posible utilizar el modelo E-R para componer música?
R. Sí, es posible. Los resultados de este trabajo de investigación así lo proponen.
- ¿Es posible desarrollar mecanismos computacionales que puedan generar una representación de las emociones en una melodía?
R. Se han generado las rutinas necesarias para construir una representación de emociones, basadas en lógica difusa.
- ¿Es posible utilizar la representación de emociones para ligar ideas musicales en lugar de teoría musical?
R. Tal representación se utiliza como guía del proceso de composición, con ella es posible enlazar ideas musicales.
- ¿Es posible generar un programa de computo que represente el proceso de composición como lo realiza alguien que no tiene profundos conocimientos musicales?
R. Es posible, ya que el proceso de composición esta guiado por niveles

de emociones y tensiones, y se ha encontrado que las salidas del sistema se evalúan con similitud a las compuestas por humanos. Para lograr esto no ha sido necesario recurrir a la teoría musical.

El sistema de composición producto de esta investigación, fue presentado en el *5th International Joint Workshop on Computational Creativity* en Madrid, España. Para mayor información se puede consultar [Alvarado & Pérez y Pérez, 2008].

Apéndice A

Entrevista

Esta entrevista es parte de una investigación sobre generación de melodías musicales. El objetivo es conocer las opiniones de personas acerca de cuatro melodías que serán presentadas. Nuestro principal interés es que usted evalúe la composición de la melodía; no estamos interesados en la evaluación de la interpretación o la instrumentación.

DATOS PERSONALES:

Edad:

Sexo:

Nivel de estudios o profesión:

¿Tiene estudios musicales? (especifique qué tipo de estudios):

INSTRUCCIONES:

- Se reproducirán de forma continua cuatro melodías en un orden aleatorio. Cada una de ellas tiene una duración aproximada entre 32 y 38 segundos y las cuatro son interpretadas empleando MIDI y una computadora.
- Luego se repetirán una a una las mismas melodías, pero en un orden diferente al paso anterior, dejando unos minutos entre cada reproducción para que responda los cuestionarios de evaluación que se anexan a este documento. A cada melodía le corresponde un cuestionario. El siguiente es un ejemplo del cuestionario.

1. ¿La melodía que acaba de escuchar le gustó?
 - a) nada
 - b) poco
 - c) más o menos
 - d) mucho
 - e) muchísimo
2. ¿La melodía escuchada fue compuesta por?
 - a) una persona que no sabe de composiciones musicales
 - b) una persona con pocos conocimientos y experiencia como compositor
 - c) un compositor amateur
 - d) un compositor profesional principiante
 - e) un compositor profesional con experiencia
3. La composición que acaba de escuchar le parece:
 - a) una copia de otras que ha escuchado
 - b) muy similar a otras que ha escuchado
 - c) diferente a otras que ha escuchado
 - d) original
 - e) muy original
4. Con sus propias palabras diga qué le pareció la melodía que acaba de escuchar (la composición le parece aburrida, repetitiva, interesante, es una basura, etc.).

R:
5. ¿Algún comentario que desee añadir?

R:

# 1 (la mejor)	<u>Melodía 3</u>
# 2	<u>Melodía 1</u>
# 3	<u>Melodía 2</u>
# 4 (la peor)	<u>Melodía 4</u>

- En la última parte de este estudio se le pide que escriba en forma de lista, de mejor a peor, el número de cada una de las melodías que escuchó. Por ejemplo:

MUCHAS GRACIAS POR SU COLABORACIÓN EN ESTE ESTUDIO

Nos interesa conocer su opinión sobre las composiciones que a continuación se presentan.

Escuche con atención la melodía número 1 y conteste el siguiente cuestionario:

1. La melodía que acaba de escuchar le gustó
 - a) nada
 - b) poco
 - c) más o menos
 - d) mucho
 - e) muchísimo

2. ¿La melodía escuchada fue compuesta por?
 - a) una persona que no sabe de composiciones musicales
 - b) una persona con pocos conocimientos y experiencia como compositor
 - c) un compositor amateur
 - d) un compositor profesional principiante
 - e) un compositor profesional con experiencia

3. La composición que acaba de escuchar le parece:
 - a) una copia de otras que ha escuchado
 - b) muy similar a otras que ha escuchado
 - c) diferente a otras que ha escuchado
 - d) original
 - e) muy original

4. Con sus propias palabras diga qué le pareció la melodía que acaba de escuchar (la composición le parece aburrida, repetitiva, interesante, es una basura, etc.).
R:

5. ¿Algún comentario que desee añadir?
R:

Escuche con atención la melodía número 2 y conteste el siguiente cuestionario:

1. La melodía que acaba de escuchar le gustó
 - a) nada
 - b) poco
 - c) más o menos
 - d) mucho
 - e) muchísimo

2. ¿La melodía escuchada fue compuesta por?
 - a) una persona que no sabe de composiciones musicales
 - b) una persona con pocos conocimientos y experiencia como compositor
 - c) un compositor amateur
 - d) un compositor profesional principiante
 - e) un compositor profesional con experiencia

3. La composición que acaba de escuchar le parece:
 - a) una copia de otras que ha escuchado
 - b) muy similar a otras que ha escuchado
 - c) diferente a otras que ha escuchado
 - d) original
 - e) muy original

4. Con sus propias palabras diga qué le pareció la melodía que acaba de escuchar (la composición le parece aburrida, repetitiva, interesante, es una basura, etc.).
R:

5. ¿Algún comentario que desee añadir?
R:

Escuche con atención la melodía número 3 y conteste el siguiente cuestionario:

1. La melodía que acaba de escuchar le gustó
 - a) nada
 - b) poco
 - c) más o menos
 - d) mucho
 - e) muchísimo

2. ¿La melodía escuchada fue compuesta por?
 - a) una persona que no sabe de composiciones musicales
 - b) una persona con pocos conocimientos y experiencia como compositor
 - c) un compositor amateur
 - d) un compositor profesional principiante
 - e) un compositor profesional con experiencia

3. La composición que acaba de escuchar le parece:
 - a) una copia de otras que ha escuchado
 - b) muy similar a otras que ha escuchado
 - c) diferente a otras que ha escuchado
 - d) original
 - e) muy original

4. Con sus propias palabras diga qué le pareció la melodía que acaba de escuchar (la composición le parece aburrida, repetitiva, interesante, es una basura, etc.).
R:

5. ¿Algún comentario que desee añadir?
R:

Escuche con atención la melodía número 4 y conteste el siguiente cuestionario:

1. La melodía que acaba de escuchar le gustó
 - a) nada
 - b) poco
 - c) más o menos
 - d) mucho
 - e) muchísimo

2. ¿La melodía escuchada fue compuesta por?
 - a) una persona que no sabe de composiciones musicales
 - b) una persona con pocos conocimientos y experiencia como compositor
 - c) un compositor amateur
 - d) un compositor profesional principiante
 - e) un compositor profesional con experiencia

3. La composición que acaba de escuchar le parece:
 - a) una copia de otras que ha escuchado
 - b) muy similar a otras que ha escuchado
 - c) diferente a otras que ha escuchado
 - d) original
 - e) muy original

4. Con sus propias palabras diga qué le pareció la melodía que acaba de escuchar (la composición le parece aburrida, repetitiva, interesante, es una basura, etc.).
R:

5. ¿Algún comentario que desee añadir?
R:

1 (la mejor) _____
2 _____
3 _____
4 (la peor) _____

En su opinión ¿cuál es el orden (de mejor a peor) de las melodías escuchadas?
Escriba en las líneas en blanco el número de la melodía.

¿Algún comentario que desee añadir sobre este estudio?

Apéndice B

Partituras

En este apéndice se muestran las partituras de la melodías utilizadas durante la entrevista descrita en el capítulo 6

Melodía 1 de ERMEG

ERMEG

The musical score for "Melodía 1 de ERMEG" is presented in a single system with seven staves. The music is written in 4/4 time and begins with a treble clef. The first staff starts at measure 1 and includes a triplet of eighth notes in measure 4. The second staff starts at measure 3. The third staff starts at measure 5 and features a dense eighth-note pattern in measure 6. The fourth staff starts at measure 8. The fifth staff starts at measure 10. The sixth staff starts at measure 12. The seventh staff starts at measure 16 and includes two triplet markings over eighth notes in measures 18 and 19. The piece concludes with a double bar line at the end of the seventh staff.

, *Figura B.1: Melodía 1 del ERMEG*

Melodía 2 de ERMEG

ERMEG

The image displays a musical score for a piece titled "Melodía 2 de ERMEG". The score is written in a single system on a grand staff (treble clef) in 4/4 time. The key signature is one sharp (F#), indicating the key of D major. The score consists of 18 measures, with measure numbers 1, 3, 6, 10, 11, 12, 14, 15, 16, and 18 explicitly labeled at the beginning of their respective lines. The melody features various rhythmic patterns, including eighth and sixteenth notes, and includes several triplet markings (indicated by a '3' above the notes). The piece concludes with a double bar line at the end of measure 18.

, *Figura B.2: Melodia 2 de ERMEG*

To Live is to Die (fragmento)

Metallica

The image displays a musical score for the song "To Live is to Die" by Metallica, specifically a fragment. The score is written in 4/4 time and consists of seven staves of music, numbered 1 through 15. The notation includes various rhythmic values such as quarter notes, eighth notes, and sixteenth notes, along with rests and accidentals. The key signature is one flat (B-flat major or D minor). The score begins with a treble clef and a 4/4 time signature. The first staff (measure 1) starts with a quarter rest, followed by a quarter note G4, a quarter note F#4, a quarter note E4, and a quarter note D4. The second staff (measures 2-4) continues the melody with eighth and quarter notes, including a quarter rest in measure 3. The third staff (measures 5-6) features a quarter note G4, a quarter note F#4, a quarter note E4, and a quarter note D4, followed by a quarter rest in measure 5. The fourth staff (measures 7-9) shows a quarter note G4, a quarter note F#4, a quarter note E4, and a quarter note D4, followed by a quarter rest in measure 7. The fifth staff (measures 10-11) contains a quarter note G4, a quarter note F#4, a quarter note E4, and a quarter note D4, followed by a quarter rest in measure 10. The sixth staff (measures 12-14) starts with a quarter rest in measure 12, followed by a quarter note G4, a quarter note F#4, a quarter note E4, and a quarter note D4. The seventh staff (measures 15-16) begins with a quarter rest in measure 15, followed by a quarter note G4, a quarter note F#4, a quarter note E4, and a quarter note D4.

, *Figura B.3: Melodia 1 de Metallica*

Fade to Black (fragmento)

Metallica

1

2

3

5

7

10

11

13

18

, *Figura B.4: Melodía 2 de Metallica*

Bibliografía

- [Alvarado & Pérez y Pérez, 2008] Alvarado López J., & Pérez y Pérez, R. (2008). A Computer Model for the Generation of Monophonic Musical Melodies. *Proc. of the International Joint Workshop on Computational Creativity*, 117-126.
- [Alvaro, Miranda & Barros, 2005] Alvaro, J. L., Miranda, E. R., & Barros, B. (2005). Music Knowledge Analysis: Towards an Efficient Representation for Composition. *Proc. The 11th Conference of the Spanish Association for Artificial Intelligence*.
- [Boden, 1998] Boden, M. (1998). Creativity and artificial intelligence. *Artificial Intelligence*, 103, 347-356.
- [Chiu & Kumar, 2003] Chiu, P., Kumar, A. (2003). Music therapy: Loud noise or soothing notes? *International Pediatrics* 18(4) 204-208
- [Colton, 2009] Colton, S. (2009). En *The painting fool*. Recuperado el 3 de Febrero de 2009, de The Painting Fool - About Me: <http://thepaintingfool.com/about/index.html>
- [“Computational creativity,” 2009] Computational creativity (2009, Enero 25) En *Wikipedia, the free encyclopedia*. Recuperado el 2 de febrero del 2009 de http://en.wikipedia.org/wiki/Artificial_Creativity
- [Cope, 1991] Cope, D. (1991). *Computers and Musical Style (Computer Music and Digital Audio Series)*. A-R Editions.
- [Dalla Bella, Peretz, Rousseau & Gosselin, 2001] Dalla Bella, S., Peretz, I., Rousseau, L., & Gosselin, N. (2001). A developmental study of the affective value of tempo and mode in music. *Cognition*, 80(3), B1-B10.

- [Gelernter, 1994] Gelernter, D. 1994. *The Muse in the Machine*. London: Fourth Estate.
- [Hofstadter, 2001] Hofstadter, D. (2001). *Staring Emmy straight in the eye- And doing my best not to flinch* (pp. 33-82). En: D. Cope (Ed) *Virtual Music: Computer Synthesis of Musical Style*. Cambridge MA and London: The MIT Press;
- [Juslin & Sloboda, 2001] Juslin, P. & Sloboda, J. A. (2001) *Music and Emotion: Theory and Research*. Oxford: Oxford University Press.
- [Khalifa, Roy, Rainville, Dalla Bella, & Peretz, 2008] Khalifa, S., Roy, M., Rainville, P., Dalla Bella, S. & Peretz, I. (2008) Role of tempo entrainment in psychophysiological differentiation of happy and sad music? *International Journal of Psychophysiology*, vol. 68, pp. 17-26
- [Mozer, 1994] Mozer, M. C. (1994). Neural Network Music Composition by Prediction: Exploring the Benefits of Psychoacoustic Constraints and Multi-Scale Processing. *Connection-Science*, 6(2-3), 247-280.
- [Newell & Simon, 1961] Newell, A., & Simon, H. A. (1961). Computer Simulation of Human Thinking. *Science*, 134(3495), 2011-2017.
- [Newell, Shaw & Simon, 1963] Newell, A., Shaw, J., & Simon, H. A. (1963), The process of creative thinking, H. E. Gruber, G. Terrell and M. Wertheimer (Eds.), *Contemporary Approaches to Creative Thinking*, pp 63 – 119. New York: Atherton
- [O'Reilly & Ortega, 2008] O'Reilly, F., & Ortega, H. (2008). En *Departamento de Probabilidad y Estadística (IIMAS, UNAM)*. Recuperado el 20 de Abril de 2008, de Juego de dados de Mozart: <http://www.dpye.iimas.unam.mx/federico/mozart/>
- [Oliveira & Cardoso, 2007] Oliveira A. P. & Cardoso A. (2007). Towards Affective-Psychophysiological Foundations for Music Production. *ACII '07: Proceedings of the 2nd international conference on Affective Computing and Intelligent Interaction*. pp. 511-522
- [Papadopoulos & Wiggins, 1998] Papadopoulos, G., & Wiggins, G. A. (1998). A Genetic Algorithm for the Generation of Jazz Melodies. *Proc of STeP'98: 8th Finnish Conference on Artificial Intelligence*.

- [Pérez y Pérez, 2007] Pérez y Pérez, R. (2007). Employing Emotions to Drive Plot Generation in a Computer-Based Storyteller. *Cognitive Systems Research*. Vol. 8, number 2, pp. 89-109.
- [Pérez y Pérez & Sharples, 2004] Pérez y Pérez, R., & Sharples, M. (2004). Three computer-based models of storytelling: BRUTUS, MINSTREL and MEXICA. *Knowledge-Based Systems*, (17), 15-29.
- [Pérez y Pérez & Sharples, 2001] Pérez y Pérez, R & Sharples, M. (2001). MEXICA: A computer model of a cognitive account of creative writing. *Journal of Experimental and Theoretical Artificial Intelligence*, 13, 119-139.
- [(Roger Wolcott Sperry, 2008)] Roger Wolcott Sperry. (2009). En *Encyclopedia Britannica*. Recuperado el 15 de Abril de 2008 de Roger Wolcott Sperry – Britannica Online Encyclopedia: <http://www.britannica.com/eb/article-9069089/Roger-Wolcott-Sperry>
- [Ross, 1995] Ross, B. J. (1995). A Process Algebra for Stochastic Music Composition. *Proc. International Computer Music Conference (ICMC '95)*, 448-451.
- [Saunders & Gero, 2002] Saunders, R., & Gero, J. (2002). How to study artificial creativity. *Proceedings of the 4th conference on Creativity & cognition*, 80-87.
- [Sharples, 1999] Sharples, M. 1999. *How We Write: Writing as Creative Design*, New York: Routledge.
- [Sharples, 1996] Sharples, M. 1996. 'An Account of Writing as Creative Design'. In C.M. Levy & S. Ransdell (1996), pp. 127-48.
- [Sharples, 1994] Sharples, M. 1994. 'Cognitive Support and the Rhythm of Design'. In T. Dartnall (1994) pp. 385-402.
- [Sharples, 1991] Sharples, M. 1991. Computational Story Writing. Cognitive Science Research Paper. School of Cognitive and Computing Science, University of Sussex, England.
- [Simon, 1989] Simon, H. A. (1989). Making management decisions: The role of intuition and emotion. *Academy of Management Executive*, 57-64.

- [Simon, 2001] Simon, H. A. (2001). Creativity in the arts and sciences. *In Cultures of creativity: The centennial celebrations of the Nobel Prizes. The Kenyon Review*, 23(2), 203-220.
- [Todd, 1989] Todd, P. M. (1989). A Connectionist Approach to Algorithmic Composition. *Computer Music Journal*, Vol. 13, No. 4 pp. 27-43, Published by: The MIT Press
- [Turner, 1994] Turner, S. R. (1994). The creative process: A computer model of storytelling. Hillsdale, NJ: Lawrence Erlbaum Associates.
- [Yu-Xi & Kim Teng, 1996] Yu-Xi, L., & Kim Teng, L. (1996). Generating music using genetic algorithms. *Proc. Science Research Congress, Singapore*, 138-144
- [Zimmermann, 1999] Zimmermann, D. (1999). Creativity - Simply Random? *Symposium on AI and Scientific Creativity (99), Edinburgh*.