

**HIDDEN MARKOV MODELS (HMM's) Y APLICACIONES**

**EDISON OSPINA MATEUS**

**UNIVERSIDAD TECNOLÓGICA DE BOLÍVAR  
FACULTAD DE INGENIERÍA ELECTRÓNICA Y ELÉCTRICA  
CARTAGENA DE INDIAS D.T Y C**

**2008**

# **HIDDEN MARKOV MODELS (HMM's) Y APLICACIONES**

EDISON OSPINA MATEUS

Monografía de Grado presentada para optar al título de Ingeniero Electrónico

**Minor en Automatización Industrial**

Director

Eduardo Gómez Vásquez

Magister en Ciencias Computacionales

UNIVERSIDAD TECNOLÓGICA DE BOLÍVAR

FACULTAD DE INGENIERÍA ELECTRÓNICA Y ELÉCTRICA

CARTAGENA DE INDIAS D.T Y C

2008

Nota de aceptación

---

---

---

---

---

---

Firma de Jurado

---

Firma de Jurado

Cartagena de Indias, D.T y C, Diciembre del 2008

Cartagena de Indias D.T y C. Diciembre 5 del 2008.

Señores

**COMITÉ FACULTAD DE INGENIERÍAS**

PROGRAMA DE INGENIERÍA ELÉCTRICA Y ELECTRÓNICA

Universidad Tecnológica de Bolívar

Cartagena

**ASUNTO: PRESENTACIÓN MONOGRAFÍA DE GRADO.**

Respetados señores:

Con toda atención me dirijo a ustedes con el fin de presentarles a su consideración, estudio y aprobación la monografía titulada "HIDDEN MARKOV MODELS (HMM's) Y APLICACIONES" como requisito parcial para optar al título de Ingeniero Electrónico.

Agradezco la atención prestada,

Cordialmente,

---

**EDISON OSPINA MATEUS**

**C.C. 73.009.161 de Cartagena**

Cartagena de Indias D.T y C. Diciembre 5 del 2008.

Señores

**COMITÉ FACULTAD DE INGENIERÍAS**

PROGRAMA DE INGENIERÍA ELÉCTRICA Y ELECTRÓNICA

Universidad Tecnológica de Bolívar

Cartagena

**ASUNTO: PRESENTACIÓN MONOGRAFÍA DE GRADO.**

Cordial saludo,

Por medio de la presente ratifico mi aceptación como director de la Monografía de Grado titulado "HIDDEN MARKOV MODELS (HMM's) Y APLICACIONES", del cual me permito hacer entrega formal del mismo, que será presentado por el estudiante del programa de Ingeniería Electrónica EDISON OSPINA MATEUS.

Agradezco la atención prestada,

Cordialmente,

---

**EDUARDO GÓMEZ VÁSQUEZ**

## **AUTORIZACIÓN**

Cartagena de Indias, Diciembre 5 de 2008.

Yo, EDISON OSPINA MATEUS, identificado con la cedula de ciudadanía número 73.009.161 de Cartagena, autoriza a la Universidad Tecnológica de Bolívar, para hacer uso de mi trabajo de monografía y publicarlo en el catalogo on-line de la biblioteca.

---

EDISON OSPINA MATEUS.

C.C. 73.009.161 de Cartagena.

## **AGRADECIMIENTOS**

El autor expresa sus mas especiales agradecimientos a:

EDUARDO GOMEZ VASQUEZ, Director de Monografía, por su confianza, constante e incondicional apoyo proporcionandome sus consejos, la experiencia y los lineamientos necesarios para que este documento fuera posible.

Todos aquellos docentes de la Universidad Tecnologica de Bolivar que durante mis estudios contribuyeron en mi formación profesional, depositando en mi sus experiencias, conocimientos, métodos, disciplinas que me forjan hoy dia como un profesional integro.

*Dedicatoria*

*A Dios, por iluminar de todas las diversas y misteriosas formas  
los caminos que conllevan a la satisfacción espiritual.*

*Por el producto de su bendición, mis padres, que  
me han criado y querido de tal manera, haciendo de mí  
un ser de buen juicio. Sin ellos este logro no hubiese sido posible.*

*A mis hermanos, Laury y Holman, por apoyarme  
y hacerme reír con sus ocurrencias cada vez que pueden.*

*A todos mis amigos por estar ahí en diversos momentos de la vida y que  
sin ellos saberlos, han contribuido en magna proporción a la  
consecución de este logro.*

*Edison*

**TABLA DE CONTENIDO**

|  | pág. |
|--|------|
| INTRODUCCIÓN   | 19   |
| 1. INTELIGENCIA ARTIFICIAL, MAQUINAS DE APRENDIZAJE, Y CADENAS OCULTAS DE MARKOV | 23   |
| 2. INTRODUCCIÓN CADENAS OCULTAS DE MARKOV  | 28   |
| 2.1. PERSPECTIVA HISTÓRICA   | 28   |
| 2.2. NATURALEZA APLICATIVA DE LAS CADENAS OCULTAS DE MARKOV                      | 29   |
| 2.2.1. Mecanismo de Aprendizaje  | 31   |
| 2.2.2. Acerca de la información que procesa                                      | 35   |
| 2.2.3. Acerca de las actividades que ejecuta                                     | 38   |
| 3. CADENAS OCULTAS DE MARKOV: ESTRUCTURA Y ANÁLISIS                              | 42   |
| 3.1. CADENAS DE MARKOV   | 42   |
| 3.2. CADENAS OCULTAS DE MARKOV (HMM)   | 49   |
| 3.2.1. Componentes de un HMM   | 50   |
| 3.3. HMM COMO GENERADOR DE OBSERVACIONES   | 53   |
| 3.4. LOS TRES PROBLEMAS BÁSICOS EN HMM's Y SUS SOLUCIONES                        | 55   |
| 3.4.1. Problema 1: Evaluación  | 55   |
| 3.4.2. Problema 2: Decodificación  | 61   |
| 3.4.3. Problema 3: Entrenamiento   | 65   |
| 3.5. TOPOLOGÍAS DE LAS HMM's   | 68   |
| 3.6. DENSIDAD DE OBSERVACIONES CONTINUÚAS EN HMM                                 | 70   |

|   |     |
|---|-----|
| 3.7. ALGUNAS DIFICULTADES DE IMPLEMENTACIÓN   | 73  |
| 4. APLICACIONES DE LAS CADENAS OCULTAS DE MARKOV  | 76  |
| 4.1. RECONOCIMIENTO AUTOMÁTICO DE VOZ   | 77  |
| 4.1.1. Naturaleza de la Generación y el Reconocimiento de Voz.  | 77  |
| 4.1.2. Selección Estructura $\lambda$ de HMM para el Reconocedor de Voz                                       | 81  |
| 4.1.3. Entrenamiento y Reconocimiento de Voz  | 87  |
| 4.2. APRENDIZAJE Y RECONOCIMIENTO DE LUGARES CON ROBOTS MÓVILES   | 90  |
| 4.2.1. Naturaleza del reconocimiento de lugares con Robots Móviles  | 90  |
| 4.2.2. Selección Estructura $\lambda$ de HMM para Robot Reconocedor de lugares                                | 94  |
| 4.2.3. Entrenamiento de los modelos y Etapa de Reconocimiento   | 102 |
| 4.3. TÉCNICAS PREDICTIVAS DE ACTIVIDADES DE MANTENIMIENTO EN LA INDUSTRIA: DIAGNOSTICO Y PRONOSTICO CON HMM'S | 108 |
| 4.3.1. Filosofía de mantenimiento de equipos industriales   | 109 |
| 4.3.2. Selección de la Estructura de un HMM para diagnostico y pronostico.                                    | 114 |
| 4.3.3. Entrenamiento y pruebas de clasificación de modelos  | 121 |
| 4.3.4. Diagnostico y Pronostico con modelos HMM   | 125 |
| 4.4. APRENDIZAJE Y RECONOCIMIENTO DE PATRONES DE MOVIMIENTO CORPORAL EN SERES HUMANOS                         | 129 |
| 4.4.1. Avances en reconocimiento de patrones de movimiento del cuerpo humano                                  | 129 |
| 4.4.2. Selección de la estructura $\lambda$ de HMM para reconocimiento de patrones de movimiento corporal     | 132 |
| 4.4.3. Entrenamiento de los modelos, Reconocimiento y Generación de movimiento                                | 139 |
| 4.4.4. Generación de movimiento a través de meta-proto-símbolos   | 141 |

|   |     |
|---|-----|
| 5. TOOLBOXES Y OTRAS HERRAMIENTAS INFORMÁTICAS EN HMM's | 147 |
| 5.1. Dhmm.tar y Hmm.tar                                 | 147 |
| 5.2. MendelHMM  | 148 |
| 5.3. UMDHMM.  | 149 |
| 5.4. HTK (Hidden Markov Model Toolkit).                 | 149 |
| 5.5. HMM_mat y HMMall                                   | 150 |
| 6. CONCLUSIONES   | 152 |

### LISTA DE FIGURAS

|   |    |
|---|----|
| Figura 1. Avances de la Inteligencia Artificial. ....                                     | 25 |
| Figura 2. Ubicación de HMM dentro de la Inteligencia Artificial. ....                     | 27 |
| Figura 3. Proceso aprendizaje Maquinas de Aprendizaje. ....                               | 31 |
| Figura 4. Espacio de parámetros para selección de probabilidades máximas. ....            | 34 |
| Figura 5. Análisis de capturas seriadas de imagen del terreno por un Robot Móvil. ....    | 37 |
| Figura 6. Problema de predicción de posibles trayectorias de una nave. ....               | 39 |
| Figura 7. Clasificación de eventos temporales desconocidos con HMM. ....                  | 40 |
| Figura 8. Proceso determinista: Iluminación del Semáforo. ....                            | 43 |
| Figura 9. Maquina de estados Cadena de Markov. ....                                       | 43 |
| Figura 10. Modelo Cadena de Markov para motor de velocidad variable ....                  | 46 |
| Figura 11. Cadena Oculta de Markov. Estados y distribución de observaciones. ....         | 50 |
| Figura 12. HMM como generador de observaciones ....                                       | 54 |
| Figura 13. Ilustración del Trellis del proceso de cálculo de la variable Forward. ....    | 59 |
| Figura 14. Proceso de cálculo de la variable Backward. ....                               | 61 |
| Figura 15. Ilustración del proceso de Viterbi para hallar la secuencia más probable. .... | 63 |
| Figura 16. Proceso de Reconstrucción (Path Backtracking) de secuencia de estados. ....    | 65 |
| Figura 17. Proceso de entrenamiento con conexión de estados. ....                         | 66 |
| Figura 18. Distintas topologías de HMM. ....  | 69 |
| Figura 19. Observaciones Continuas y distribución de probabilidad Gausiana. ....          | 71 |
| Figura 20. Mixturas Gaussianas para la densidad de probabilidad de un estado. ....        | 72 |
| Figura 21. Estructura General de un Reconocedor Automático de Voz. ....                   | 78 |
| Figura 22. Proceso de pronunciación, transmisión y recepción de la voz. ....              | 78 |
| Figura 23. Reconocimiento de un mismo texto para diferentes locutores. ....               | 79 |
| Figura 24. Identificación de Doble Proceso Estocástico en Generación de Voz. ....         | 80 |
| Figura 25. Reconocedor de palabra aislada de W elementos. ....                            | 82 |

|   |     |
|---|-----|
| Figura 26. Topología Left-to-Right de HMM utilizada en el reconocimiento de voz. ....                 | 84  |
| Figura 27. Obtención conjunto de observaciones O por pre-procesamiento de señal. .                    | 85  |
| Figura 28. Función de densidad de probabilidad de las observaciones O por Estado. .                   | 87  |
| Figura 29. Diagrama de bloques de un reconocedor de palabras aisladas con HMM. .                      | 88  |
| Figura 30. Niveles del Proceso de Navegación en Robots Móviles. ....                                  | 91  |
| Figura 31. Robots móviles para reconocimiento de lugares interiores y exteriores. ....                | 95  |
| Figura 32. Set de características del terreno en ambientes interiores y exteriores. ....              | 97  |
| Figura 33. Evolución temporal de características y Medidas de los sensores. ....                      | 98  |
| Figura 34. Topología de estados y evolución temporal de las señales de sensores. .                    | 100 |
| Figura 35. Funciones de distribución de probabilidad por estado Intersección T. ....                  | 102 |
| Figura 36. Secuencia combinada de varios lugares en ambientes interiores. ....                        | 102 |
| Figura 37. Segmentación y Etiqueta de los lugares en ambientes interiores. ....                       | 103 |
| Figura 38. Segmentación y Etiqueta de Situaciones en Ambientes Exteriores. ....                       | 104 |
| Figura 39. Reconocimiento lugares compuestos con algoritmo de Viterbi. ....                           | 106 |
| Figura 40. Selección de costos óptimos en mantenimiento de equipos. ....                              | 110 |
| Figura 41. Estructura de Mantenimiento Basado en Condición. ....                                      | 112 |
| Figura 42. Proceso de deterioro de equipos en el tiempo. ....   | 113 |
| Figura 43. Modelos de Helicopteros fabricante Westland. ....  | 114 |
| Figura 44. Sistema de transmisión de un Helicóptero Westland. ....                                    | 115 |
| Figura 45. Condición del sistema sin defectos a 27% y 100% del torque nominal. ....                   | 117 |
| Figura 46. Condición del sistema al 100% del torque nominal, con y sin defectos. ....                 | 118 |
| Figura 47. Condición del sistema a diferentes niveles de severidad a un mismo torque y defecto. ....  | 118 |
| Figura 48. Topología de estados del sistema de trasmisión del Westland para un solo defecto. ....     | 119 |
| Figura 49. Topología de estados del sistema de transmisión del Westland para todos los defectos. .... | 120 |
| Figura 50. Clasificación de condiciones de operación. ....  | 122 |
| Figura 51. Secuencia de datos para el sensor # 7. ....  | 123 |
| Figura 52. Identificación secuencia de condiciones con diferentes números de muestras. ....           | 124 |

|  |     |
|--|-----|
| Figura 53. Topología de estados para el pronóstico del sistema.....              | 126 |
| Figura 54. Reconocimiento de patrones de movimiento corporal.....                | 130 |
| Figura 55. Ejecución de tareas por método de imitación.....                      | 131 |
| Figura 56. Escenario de KickBoxing entre humano y robot.....                     | 133 |
| Figura 57. Modelo de comunicación mimética.....                                  | 134 |
| Figura 58. Representación de patrones de movimiento (protosimbolos) con HMM....  | 135 |
| Figura 59. Espacio de protosimbolos.....   | 137 |
| Figura 60. Representación de patrones de Interacción (meta-proto-símbolos) ..... | 138 |
| Figura 61. Trayectoria de símbolos resultantes de una interacción.....           | 140 |
| Figura 62. Fase de Generación de movimiento para un robot.....                   | 141 |
| Figura 63. Trayectorias de símbolos hipotéticos.....                             | 142 |
| Figura 64. Generación de movimiento por superposición de protosimbolos.....      | 143 |
| Figura 65. Capturas virtuales de la reacción del robot a estímulos.....          | 144 |
| Figura 66. Capturas en tiempo real de la reacción del robot a estímulos.....     | 145 |
| Figura 67. Ejecución del HMM.tar .....   | 148 |
| Figura 68. Imágenes de interfaz grafica del MendelHMM toolbox.....               | 149 |
| Figura 69. Estructura del programa Learn_mhmm_demo.m .....                       | 151 |

### GLOSARIO

**ANN:** sistema de inteligencia artificial conexionista inspirado en la estructura biológica de las neuronas para la ejecución de actividades que requieran de inteligencia. Conocida por su efectividad en tareas de reconocimiento y clasificación de patrones. Corresponde a las siglas en inglés (Artificial Neural Networks).

**Aprendizaje inductivo:** construcción de modelos o teorías a partir de datos recolectados. Proceso de generalizar a partir de elementos específicos.

**Clusterización:** identificación de clases dentro de un conjunto masivo de datos. Detección de irregularidades o discrepancias entre la información.

**CEPSTRUM:** resultado de tomar la transformada de Fourier (FT) del espectro de una señal. Su nombre se deriva de revertir las primeras cuatro letras "spectrum".

**Determinista, proceso:** sistema en el cual el azar no está involucrado en los futuros estados del sistema. Es decir, si se conoce el estado actual del sistema, las variables de ambiente y el comportamiento del sistema ante los cambios en el ambiente, entonces se puede predecir sin ningún riesgo de error el estado siguiente del sistema.

**Ergodicidad:** transformación que preserva la medida  $T$  en un espacio de probabilidad, se dice que es ergódico si un conjunto medible que es invariante bajo  $T$ , tiene medida 0 ó 1. Un antiguo término para esta propiedad era métricamente transitivo.

**Espacio de hipótesis:** conjunto de soluciones factibles en la asignación de algunas variables paramétricas que satisfagan la conclusión de algunas premisas de un problema específico.

**Estocástico, proceso:** es un concepto matemático que sirve para caracterizar y estudiar todo tipo fenómenos aleatorios (estocásticos) que evolucionan, generalmente, con el tiempo.

**Feature extraction:** reducción dimensional de los datos de entrada cuando esta es redundante, y no contiene información relevante.

**Iteración:** proceso cíclico de mejoramiento continuo para la consecución de parámetros óptimos en algún modelo.

**Kalman, filtros de:** algoritmo desarrollado por Rudolf E. Kalman en 1960 que sirve para poder estimar el estado oculto (no medible) de un sistema dinámico lineal.

**LISP:** es el segundo lenguaje de programación, después de Fortran, de alto nivel. Lisp es de tipo declarativo y fue creado en 1958 por John McCarthy y sus colaboradores en el MIT.

**LPC:** (Linear predicting coding). Herramienta usada por lo general en el procesamiento de voz, para representar envolventes espectrales de la señal con un modelo predictivo.

**Patrón:** características particulares de una variable, evento o proceso tomadas como referencia por ser únicas en su tipo. A partir de un patrón desciende características con ligeras variaciones consideradas como hijos.

**PCA:** (Principal Component Analysis) es un transformador de vectores espaciales utilizado para reducir conjuntos de datos multidimensionales a un conjunto de dimensión reducida para el análisis.

**Perceptrón:** es un tipo de red neuronal artificial desarrollado por Frank Rosenblatt, también puede entenderse como perceptrón la neurona artificial y unidad básica de inferencia en forma de discriminador lineal, que constituye este modelo de red neuronal

artificial, esto debido a que el perceptrón puede usarse como neurona dentro de un perceptrón mas grande u otro tipo de red neuronal artificial.

**Pitch:** representa la frecuencia fundamental de un sonido percibido.

**Prolog:** es un lenguaje de programación lógico e interpretado, bastante popular en el medio de investigación en Inteligencia Artificial. Proveniente del francés Programation et Logique

**Toolbox:** palabra equivalente a bloque de herramientas en ingles, que contiene un software o algoritmo diseñado específicamente para la solución de algún tipo un tipo de problema dentro de una área, ciencia, concepto, etc.

**Topología:** forma en que la estructura de un modelo interconecta sus elementos entre sí para manifiestar un grado de relación.

**SVM:** maquina de aprendizaje llamado Maquina de Soporte Vectorial, que es conocido por su efectividad en aplicaciones que requieren de clasificación de datos. Corresponde a las siglas en ingles Support Vectorial Machine.

**Viterbi:** El algoritmo de Viterbi permite encontrar la secuencia de estados más probable en un Modelo oculto de Márkov.

### RESUMEN

La intención principal de este trabajo de monografía es dar a conocer una maquina de aprendizaje existente desde hace algunos años , conocida como Cadenas Ocultas de Markov (HMM: Hidden Markov Models), mostrando su estructura matemática, componentes y naturaleza aplicativa con el fin de conocer el tipo de información, el método de procesamiento y las tareas que puede resolver. Como atractivo principal del documento, está el de identificar los elementos teóricos de un HMM dentro de un conjunto de aplicaciones reales ya implementadas por algunos científicos. Se describen por lo tanto 4 aplicaciones de diferentes naturalezas para visualizar el alcance de implementación de estas maquinas. Entre las aplicaciones seleccionadas están; el reconocimiento de voz, reconocimiento de lugares con robots móviles, técnicas predictivas para el mantenimiento de equipos y por último, aprendizaje y reconocimiento de patrones de movimiento corporal en seres humanos como método de interacción hombre-máquina. Se hace un acercamiento a las herramientas informáticas disponibles en la WEB que poseen algoritmos de entrenamiento, generación y prueba en HMM's. Entre ellos se muestra como ejemplo general, un diagrama del procedimiento de entrenamiento de una HMM para observaciones continuas compuestos de múltiples mixturas gaussianas de un algoritmo contenido en el toolbox de MATLAB, HMM\_mat.

### INTRODUCCIÓN

Con los avances tecnológicos que se han presentado en el campo de las ciencias computacionales, se espera que múltiples problemas que requieren ser resueltos por métodos que imiten ciertos aspectos de la inteligencia humana sigan siendo impulsados bajo este marco creciente de desarrollo. Estos avances han permitido en la actualidad solucionar problemas en amplio rango de aplicaciones de las ciencias de forma rápida y eficiente, mejorando el desempeño de los sistemas que procesan información y ejecutando tareas en áreas de control automático, robótica, economía, biología, finanzas, entretenimiento, entre otras.

Procurando fortalecer los canales de interacción Hombre-Máquina-Entorno y de acelerar la comprensión del comportamiento de los sistemas y fenómenos naturales que se presentan, durante años los científicos se habían propuesto previamente cuestiones como las siguientes:

- ¿Es posible aprovechar esos avances tecnológicos que se encuentran en nuestro alcance para proporcionarles a las máquinas cierto grado de inteligencia, y así soporten en la ejecución de algunas actividades riesgosas que hoy en día solo seres humanos pueden llevar a cabo? ¿Podría realmente considerarse a estas máquinas como seres inteligentes? ¿Cómo cuantificaríamos esta cualidad?
- ¿Sería posible darle facultades a un sistema para organizar, clasificar, reconocer, tomar decisiones y diagnosticar a partir de información tomada del entorno como lo hacen los seres humanos? ¿Podría estos sistemas **adaptarse** autónomamente al entorno, si este cambiase?

Muchos años atrás, se pensaba que estas ideas eran producto de la ciencia ficción, y sin embargo en la actualidad hemos sido testigos de los intentos y los resultados obtenidos por los científicos por crear entidades que traten hasta cierto nivel tener “algo” de inteligencia creando a mediados de los 50's el conocido movimiento de la Inteligencia Artificial.

Intuitivamente surgiría una nueva pregunta, ¿Que se necesita para desarrollar un sistema inteligente?. El diseño de un sistema inerte que imite algunos comportamientos que desarrollan los seres humanos que naturalmente a estos últimos les resulta una tarea de simplicidad instintiva o algunas veces entrenada, es una tarea evidentemente compleja que requiere de análisis de muchos aspectos. Como consecuencia, se haría esfuerzos en proporcionarle una estructura “cerebral” o de procesamiento, y unos mecanismos autónomos de interacción con el entorno con el fin de que el sistema “aprenda” conocimientos a partir de la información que percibe de su proceso de interacción. El conocimiento obtenido sería pues el resultado de tal proceso de aprendizaje, representado con un modelo simbólico que describe el proceso ejecutado en la interacción.

En la mayoría de los casos, describir o modelar tal clase de procesos inteligentes excluye el uso de métodos deterministas (reglas lógicas o ecuaciones), y las programaciones manuales, que en contraste si se requeriría en la computación tradicional. La causa principal de no usarlos, recae en la exigencia que estos métodos imponen a la información a aspectos como la tolerancia a la variabilidad, a la capacidad de abstraer información útil a partir de grandes cantidad de datos en corto tiempo (compresión) y a algunas incertidumbres que podría aparecer dentro de ella. La presencia de estos factores en la información a analizar (datos anómalos, incompletos, distorsionados, excesivos o incomprensibles) para sistemas de computación tradicional lo conduce a afecciones en su desempeño y en el peor de los casos el colapso.

Como respuesta a la necesidad de modelar procesos inteligentes, se han estado desarrollando dentro de un sub-campo de la Inteligencia Artificial, las máquinas de aprendizaje (*Machine Learning*), los cuales dan un tratamiento especial a la información con problemas de incertidumbre, variabilidad, dinamismo, y cantidad con métodos computacionales y estadísticos. Estos se enfocan en el diseño y desarrollo de modelos algorítmicos que reproduzcan dentro de un margen aceptable un proceso real de interés en forma simplificada en base a la información relevante adquirida en una etapa de entrenamiento con las características mencionadas.

Múltiples máquinas de aprendizaje se han estado desarrollando, entre ellos Árboles de decisión de aprendizaje, las Maquinas de Soporte Vectorial (SVM), Redes Bayesianas, Cadenas Ocultas de Markov (HMM's), e hibridaciones con ciencias bio-inspiradas como las Redes Neuronales Artificiales (ANN's), los Algoritmos Genéticos entre otras.

El objetivo de este documento es dar especial atención a las Cadenas Ocultas de Markov, o HMM's (Hidden Markov Models), considerado como una de las maquinas de aprendizaje con gran desempeño y robustez en un amplio rango de aplicaciones como monitoreo y control de procesos industriales, diagnóstico médico e industrial, interpretación de imágenes de video, robótica, reconocimiento de voz, detección de fallas en sistemas dinámicos, biología molecular, aeronáutica, psicología, neurociencia, sistemas de comunicación, entre otras.

La herramienta utilizada por HMM para analizar la información u observaciones que percibe del entorno está fundamentada en un lenguaje: Teoría de Probabilidades. HMM se fundamenta en procesos doblemente estocásticos para modelar con la mayor probabilidad un proceso o tarea particular que se desarrolla en el tiempo, identificando los patrones secuenciales que lo rige. A partir de dicho modelo se analiza nuevos patrones desconocidos que caracterizan a nueva información captada (por sensores o base de datos) de un proceso real, permitiéndoles al HMM entrenado tener el criterio de clasificar a este nuevo patrón dentro de una categoría establecida, estipular una secuencia de subprocesos (estados) dinámicos dentro del proceso, hacer una identificación de un evento novedoso, una predicción de una condición futura, reconocimiento de estado actual del proceso, entre otras tareas.

Este documento dará a conocer en una forma didáctica la naturaleza aplicativa de las HMM's, sus componentes principales, los problemas a los cuales se enfrentarían estas maquinas y algunos puntos importantes de implementación. Adicionalmente, se mostraran varias aplicaciones desarrolladas por científicos conocedores de la materia haciendo énfasis en la identificación de los componentes principales de las HMM's implementadas en cada una de las aplicaciones en particular. Se seleccionó aplicaciones de diversas áreas para visualizar el amplio rango de implementación de esta máquina.

Este documento se desarrollara de la siguiente forma:

En el capítulo 1 se contextualizará las Cadenas Ocultas de Markov dentro del las corrientes de pensamiento de las Inteligencia artificial en su recorrido histórico.

En el capítulo 2 se mostrara un poco los antecedentes que influyeron en el desarrollo de las HMM's y su naturaleza aplicativa vista desde un esquema "Entrada-Procesamiento – Salida" para conocer el tipo de información que acepta esta máquina, la manera en que procesa la información y aprende a partir de ella, y finalmente identificar las tareas que podrían ejecutarse con el aprendizaje obtenido.

En el capítulo 3 se presentará una breve descripción de los componentes principales para el diseño de una HMM, sus propiedades, topología y presentación formal de solución a los 3 problemas planteados en la aplicación de las HMM's.

En el capítulo 4 se presentaran 4 aplicaciones desarrolladas con HMM's por expertos en diferentes campos de la ciencia, mencionando entre ellos el reconocimiento de voz, el reconocimiento de lugares con robots móviles, técnicas predictivas de mantenimiento para el diagnostico y pronostico de condiciones de operación de equipos, y finalmente, el reconocimiento de patrones de movimiento corporal en seres humanos en los cuales se presentará la interacción de un robot humanoide y un ser humano. Se resaltaran en forma particular para cada aplicación, la identificación de cada uno de los principales componentes de un HMM, su entrenamiento y pruebas en casos reales.

En el capítulo 5 se describirá en forma breve algunos programas informáticos obtenidos de la WEB dedicados a demostrar por medio de sus algoritmos la implementación de HMM's.

Finalmente en el capítulo 6 será dedicado a las conclusiones de la investigación.

### 1. INTELIGENCIA ARTIFICIAL, MAQUINAS DE APRENDIZAJE, Y CADENAS OCULTAS DE MARKOV

Hace más de 50 años se dio inicio el desarrollo de la inteligencia artificial (IA) en donde su motivación principal era el de crear programas para máquinas que imiten el comportamiento y la comprensión humana. La investigación en el campo de la IA se caracteriza por la producción de máquinas para la automatización de tareas que requieran un comportamiento inteligente<sup>1</sup>.

Sus implementaciones en diferentes áreas de la ciencia proporcionan a los seres humanos de facilidad de afrontar problemas que contribuyan a reducir riesgos de accidentalidad al maniobrar ciertos procesos, tomar decisiones correctamente dentro de estos en base a historiales (experiencia), comprender su desarrollo en función del tiempo, modelar algunos procesos biológicos, resolver algunas tareas complejas como lo haría un experto en un área específica, etc., proporcionando mejoras en la calidad de vida, y en el desempeño de ciertos procesos industriales.

Alan Turing, con su trabajo, despertó en los científicos la idea del IA en donde planteaba un test que prueba las capacidades de una máquina para entablar una interacción con un ser humano. Entre los años 50's y 80's las investigaciones no se hicieron esperar: Jonh McCarthy y Alain Colmerauer inventaron los lenguajes de programación de alto nivel LISP y Prolog, respectivamente, para declaraciones lógicas de acciones inteligentes. Josehp Weisenbaum construyó el ELIZA, un programa

---

<sup>1</sup> Inteligencia Artificial: [http://es.wikipedia.org/wiki/Inteligencia\\_artificial](http://es.wikipedia.org/wiki/Inteligencia_artificial)

interactivo que sigue una conversación en lenguaje Ingles de cualquier tema. Edwar Feigenbaum inicia DENDRAL, una investigación de 10 años en el desarrollo de software para deducir la estructura molecular de componentes orgánicos usando datos de instrumentación científica, considerándose a este trabajo como el primer Sistema Experto. La investigación de estos sistemas se vieron reforzados gracias al trabajo de Ted Shortliffe con la creación del programa MYCIN que demostraba una aproximación basada en reglas muy practica para el diagnostico medico incluso con presencia de incertidumbre.<sup>2</sup>

Sin embargo los avances más notorios se vieron desde los años 90's (Figura 1). En 1994, Ernest Dickmanns y Daimles-Benz desarrollan los carros robóticos VITA-2 y VamP el cual transportan pasajeros a más de 1000 kilómetros en una autopista de Paris de 3 carriles con tráfico pesado a una velocidad de 130 km/h. Demostraron que el sistema se manejaba autónomamente en carriles libres, con otros automóviles, y cambios de carril automáticamente. En 1997, IBM crea el Deep Blue, una maquina de jugar ajedrez que derrotó al campeón mundial Garry Karparov. En 1999, la compañía SONY introduce al mercado el AIBO, una mascota de inteligencia artificial. En el 2004, DARPA inaugura el DARPA Gran Challenge en la cual requiere competidores para producir vehículos autónomos. Hiroaki Kitano ha formulado la Federación Internacional RoboCup, un evento en el que Robots humanoides juegan un partido de futbol para motivar aun más la investigación en IA. Muy recientemente la compañía HONDA mostró al público un robot humanoide llamado ASIMO con una gran habilidad motriz de modo casi natural.

Es así, que un sin número de motivaciones impulsan la investigación en IA hoy día de manera impresionante en diferentes aéreas de la ciencia.

---

<sup>2</sup> RICH, Elaine. Inteligencia Artificial. 2 ed. Madrid: McGraw-Hill, 1994. 703p. ISBN 84-841-1858-8.



**Figura 1. Avances de la Inteligencia Artificial<sup>3</sup>.**

En todo el trayecto investigativo de la inteligencia artificial, se le ha podido clasificar en dos grandes corrientes o escuelas de pensamiento: Inteligencia Computacional e Inteligencia Artificial Convencional. El primero envuelve desarrollos iterativos o aprendizaje por ajuste de parámetros, como los sistemas conexionista (o bio-inspirados). El aprendizaje es basado en datos empíricos y se asocia con nombres como IA no simbólica, IA desorganizada (Scruffy) o Soft-Computing. Se destacan en este pensamiento las Redes Neuronales, los sistemas Difusos (Fuzzy), la computación evolutiva inspirados en conceptos biológicos la mutación, población y supervivencia del más fuerte que se reflejan en técnicas como los algoritmos evolutivos (genéticos), inteligencia de colonias, entre otras.

La segunda corriente utiliza métodos formales y estadísticos en los cuales se enfocan en imitar la inteligencia humana a través de la manipulación de símbolos y bases de

---

<sup>3</sup> Imágenes tomadas de: [www.wikipedia.org](http://www.wikipedia.org). Palabra clave: Inteligencia artificial.

conocimiento estructuradas simbólicamente. Se le ha catalogado además a esta corriente como una técnica organizada (Neats) por el uso de la lógica y la representación formal del conocimiento. Se destacan en este campo los Sistemas Expertos, Las Redes Bayesianas, Razonamiento basado en casos, e IA basada en comportamientos.

A esta ultima corriente se le ha catalogado el nombre de **MAQUINAS DE APRENDIZAJE**, y han tenido fuerza durante los 90's y 2000's siendo muy influenciada por la teoría de la probabilidad y la estadística. Las Redes Bayesianas han sido las pioneras de este movimiento, proveyendo bases conceptuales y enlaces muy fuertes a otros temas en ingeniería como las Cadenas de Markov, Cadenas Ocultas de Markov (HMM: Hidden Markov Models), los Filtros de Kalman, Redes (Arboles) de Decisión, procesos de Poisson, procesos Gaussianos, entre otras

La maquinas de aprendizaje han fundamentado su aplicación en diferentes nichos<sup>4</sup>: 1) La minería de datos usando datos históricos para mejorar la toma de decisiones en diagnostico y pronostico de eventos; 2) Los programas o software que no pueden ser programados a mano, como el reconocimiento de patrones de voz, el manejo automático de un automóvil, el reconocimiento de un lugar, etc.; 3) Los programas que se ajustan a los peticiones, gustos y costumbres del usuario, como los sitios WEB que aprenden los intereses informativos de sus clientes, o el aprendizaje de comportamiento a través de la demostración, etc.

El funcionamiento de las maquinas de aprendizaje se basan en el concepto de mejorar la ejecución de una tarea con respecto a una medida de rendimiento basado en la experiencia, con un numero de ejemplos de entrenamiento de una situación específica. La mayoría de problemas a resolver requiere que estas maquinas se sometan a un proceso de aprendizaje similar a los que afronta un ser humano para que adquieran capacidades y habilidades particulares ante diferentes esquemas del entorno.

---

<sup>4</sup> MITCHELL, Tom. Machine Learning. s.l.: McGraw Hill, 1997. p.2-19.

## Hidden Markov Models (HMM's) y aplicaciones

Particularmente, las Redes Bayesianas, fundamentado bajo el teorema de probabilidad, adquiere un aprendizaje con el ajuste automático de sus parámetros probabilísticos que represente la situación particular a modelar. Consecuentemente, las Cadena Ocultas de Markov, como un caso particular de Red Bayesiana, funciona bajo el mismo fundamento, caracterizándose este por el modelamiento de eventos aleatorios que ocurren en el tiempo a través de información de origen ruidoso. En la sección 2.2 se mencionara la naturaleza aplicativa de los HMM's el cual enmarca los tipos de problemas que puede solucionar.

En la Figura 2 se resume la ubicación de las Cadenas Ocultas de Markov dentro de la inteligencia artificial, mencionando solo algunos representantes de cada corriente.

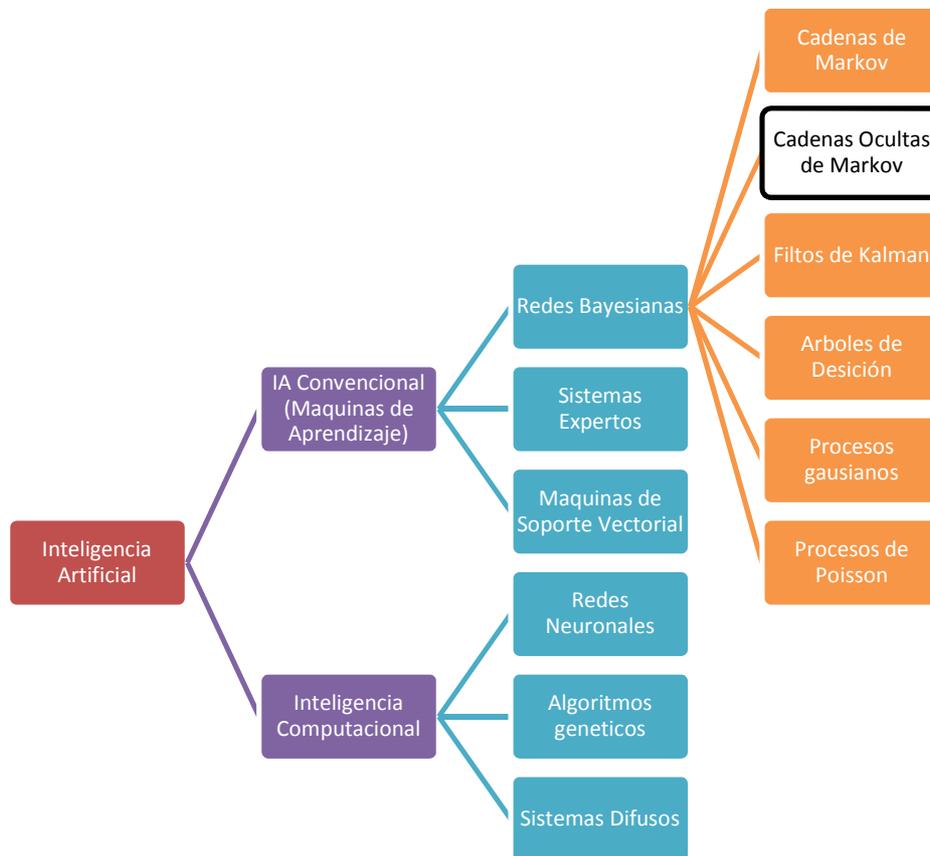


Figura 2. Ubicación de HMM dentro de la Inteligencia Artificial.

## 2. INTRODUCCIÓN CADENAS OCULTAS DE MARKOV

### 2.1. PERSPECTIVA HISTÓRICA

Las cadenas ocultas de Markov (HMM) fueron desarrolladas a principios de los años 70's y han tenido gran impacto debido a la riqueza en su estructura matemática que proporciona flexibilidad, adaptabilidad, eficiencia computacional y facilidad para modelar procesos con variaciones dinámica y temporales.

Sus fundamentos se basan en las Cadenas de Markov o procesos markovianos desarrollados mucho antes por Andréi Andreyevich Markov en el año de 1907, definiendo una maquina con un número de estados finitos en tiempo discreto. Sin embargo, este concepto es sustentado gracias a los trabajos de Thomas Bayes relativos a su Teoría de Probabilidad en 1764 y al teorema que hace honor a su nombre que es utilizado para conocer la probabilidad a posteriori (probabilidad condicional) de cierta variable de interés dado un conjunto de hallazgos.

La teoría básica de las Cadenas Ocultas de Markov fue publicada por primera vez en una serie de artículos científicos por Leonard E. Baum y sus colegas a finales de los años 60's y principios de los 70's y fue implementado para aplicaciones de procesamiento de voz por J.K Baker en Carnegie Mellon University y por Frederick Jelinek y sus colegas en IBM en los años 70's. Sin embargo, el estudio y masiva aplicación de las HMM's en el procesamiento de voz ocurrió solo hasta mediados de los 80's, debido a varias razones. Una de ellas es que la teoría básica de las HMM's fue publicada en magazines de matemáticas en los que rara vez eran leídos por ingenieros que trabajaban en procesamiento de voz. Una segunda razón se debe a que las aplicaciones originales de la teoría del procesamiento de voz no aportaban tutorial suficiente para la mayoría de los lectores que necesitaban entender la teoría y

aplicarla a sus propias investigaciones. Como resultados, varios artículos científicos fueron escritos en los cuales proveían suficiente nivel de detalle para un gran número de investigaciones que comenzaban a trabajar usando HMM en aplicaciones de procesamiento de voz.

Hoy día la aplicación de HMM no se limita a reconocimiento de voz, si no que áreas de investigación como la robótica móvil, ingeniería genética, finanzas, mantenimiento predictivo industrial, sicología, nanotecnología, comunicaciones, entre otras, que hacen uso de las herramientas matemáticas de esta metodología para analizar eventos que ocurren en forma aleatoria a través del tiempo. Todas estas aplicaciones son definidas como procesos doblemente estocásticos en el que determinaba una secuencia de variables aleatorias en el cual la variable futura es condicionada con la variable presente, pero independiente de la manera en que el estado presente haya sido condicionado por sus predecesores.

En la siguiente sección, se explicara la naturaleza de un HMM cuyo conocimiento es útil para determinar si existe la posibilidad de implementarla en aplicaciones que se deseen desarrollar. Conocer su naturaleza implica analizar qué tipos de datos o información manipula (entradas), que mecanismo de aprendizaje adopta (Procesamiento de información) y lo más importante, que actividades puede ejecutar acorde a lo aprendido en el procesamiento (Salidas).

### **2.2. NATURALEZA APLICATIVA DE LAS CADENAS OCULTAS DE MARKOV**

Desde una perspectiva general, la diversidad en maquinas de aprendizaje surge como solución a diferentes tipos de aplicaciones específicas. La naturaleza de la aplicación y las tareas a ejecutar condicionan la selección de la máquina de aprendizaje que mejor solución provea. Es por tanto, que algunas maquinas son mejores que otras por ejemplo, para clasificar un objeto, decodificar comportamientos, reconocer o identificar algún patrón, predecir eventos en un sistema, etc. por la adaptación de su filosofía al problema dado.

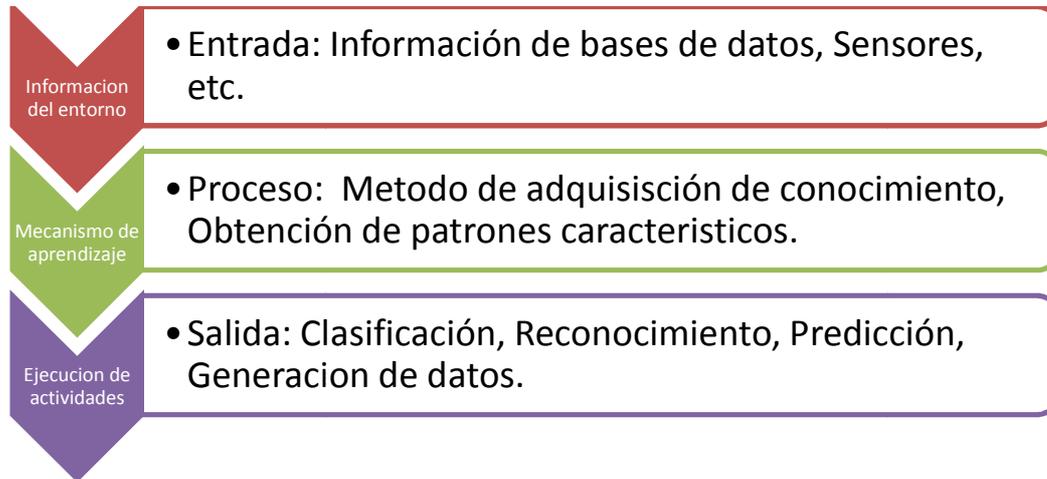
El objetivo de esta sección es conocer la naturaleza aplicativa de un HMM para determinar qué tipo de problemas es posible su implementación.

Conociendo la variedad de maquinas de aprendizaje con diferentes intenciones, objetivos y contextos de aplicación, se necesita saber cuál aplicar cuando un problema específico que requiera de inteligencia sea presentado. Es necesario por lo tanto, entender la naturaleza del problema y enmarcar el contexto de aplicación para seleccionar alguna técnica compatible con ella. Entonces se plantean las siguientes cuestiones:

- *¿Puede la estructura matemática de un HMM enmarcarse a la naturaleza de la aplicación? ¿Puede el HMM ejecutar la tarea requerida?*
- *¿A qué tipo de información se enfrenta el HMM? ¿Qué manejo le da a esa información? ¿Cómo se evalúa su desempeño en el aprendizaje?*

Acorde a estos cuestionamientos, se deben tener en cuenta los siguientes aspectos de la maquina seleccionada para la aplicación: 1) Mecanismo o método de aprendizaje por el cual la información estará sometida; 2) Tipo y características de la información o evidencia que recibirá la maquina; 3) Tareas que puede ser desarrolladas por la maquina después del aprendizaje.

En la **¡Error! No se encuentra el origen de la referencia.** se muestra un esquema que enmarca los aspectos anteriores (similar al proceso de aprendizaje de un ser humano) donde se plasma el esquema Entrada - Proceso - Salida.



**Figura 3. Proceso aprendizaje Maquinas de Aprendizaje.**

A continuación se analizan estos aspectos en forma generalizada en las Cadenas Ocultas de Markov.



**2.2.1. Mecanismo de Aprendizaje.** Existen diferentes visiones con respecto al concepto de aprendizaje aplicado a maquinas inteligentes.

Una visión genérica<sup>5</sup> establece que el aprendizaje es el proceso de mejorar el comportamiento a partir de la experiencia. Una visión más estática define como la *identificación de patrones*, de *regularidades*, existentes en la evidencia. Una visión externa se establece como la *predicción* de observaciones futuras con *plausibilidad*. Una visión teórico-informacional, (Solomonoff 1966) se resume como la *eliminación de redundancia* como un método de *comprimir la información*.

Es claro que la adquisición de un conocimiento está direccionada a la intención u objetivo del proceso de aprendizaje de una actividad específica, sea hacer una clasificación de alguna situación u objeto, tomar una decisión con base a hechos pasados, predecir un evento de cualquier índole científica, o hacer un reconocimiento de algún evento.

---

<sup>5</sup> Ibid., p.2-19.

Un método naturalmente efectivo para la adquisición de conocimiento, es someter al ente artificial al ambiente con el que interactuará con el fin de que extraiga automáticamente el conocimiento necesario a partir de las evidencias u observaciones que el contexto presente. Entonces, ¿Como una maquina representa el conocimiento adquirido durante la interacción? ¿Cómo abstrae la información necesaria?. La mayoría de las maquinas lo hacen ajustando un grupo de parámetros definidos internamente que representan en cierta forma el conocimiento obtenido. Como resultado, la obtención de valores óptimos de los parámetros dentro de una maquina en su proceso de aprendizaje, es un buen indicativo que se ha adquirido un conocimiento confiable del entorno o proceso con el cual interactuó.

Para mejorar el proceso de aprendizaje, una interacción repetitiva del sistema con diferentes condiciones del entorno es requerida como parte de un proceso de **Entrenamiento** (como lo haría un ser humano), cuyo fin es la identificación de información básica y única de una situación particular del contexto que será denominado como **Patrón**. El objetivo de reconocer patrones es clasificar objetos del entorno dentro de un número de clases o categorías. Estos objetos pueden ser imágenes, señales en formas de onda o cualquier tipo de medida que tiene que ser clasificado, siempre y cuando sea referido como patrones. El reconocimiento de patrones se ha convertido en un componente importante de los sistemas inteligentes y es usado para el procesamiento de datos y la toma de decisiones<sup>6</sup>.

Dependiendo de la cantidad y el tipo de conocimiento disponible al sistema antes de la etapa de aprendizaje (conocimiento a priori del sistema) se puede distinguir varias situaciones<sup>7</sup>:

- La forma más simple de aprendizaje se hace asignando valores a los parámetros especificados. Esta es una situación cuando el sistema contiene todo el conocimiento requerido para un tipo particular de tareas.
- El proceso para la adquisición de conocimiento en un sistema es un tipo de aprendizaje donde algunas estructuras predefinidas (reglas, esquemas etc.) son

---

<sup>6</sup> MANTYLA, Vesa-Matti. Discrete hidden Markov Models with application to isolated user-dependent hand gesture. Technical Research Centre of Finland : VTT publications 449. p.9-11.

<sup>7</sup> MARKOV, Zdravko. Machine Learning Course. May 21, 1998. University of Sofia. Faculty of Mathematics and Informatics. p.5-6.

diligenciadas con datos específicos directa o indirectamente con un experto. En este caso solo la estructura del conocimiento es conocida.

- Al sistema se le proporciona un conjunto de ejemplos (datos de entrenamiento) para que este crea una descripción de este conjunto en términos de un lenguaje particular. El conocimiento previo del sistema es la sintaxis del lenguaje permitido y posiblemente algunas características del dominio de donde proviene los ejemplos (dominio de conocimiento). Esto es una tarea típica de aprendizaje inductivo o formación de concepto.

Al igual que el aprendizaje humano, el proceso de las maquinas de aprendizaje es afectado por la presencia (o ausencia) de un profesor. En los sistemas de *aprendizaje supervisado* el profesor especifica explícitamente la salida del sistema (por ejemplo, la clase o el concepto) cuando un ejemplo es presentado al sistema (por ejemplo, los sistemas que usan datos preclasificados). En los sistemas de aprendizaje reforzado la salida exacta del sistema es desconocida, y escasamente un estimativo de su calidad es usado para guiar el proceso de aprendizaje. Por otro lado, la clusterizacion de conceptos (formación de clústeres) es un ejemplo de *aprendizaje no supervisado*. El objetivo de tales sistemas es el de analizar los datos y clasificarlos en categorías o encontrar algunas regularidades interesantes en los datos sin ejemplos de entrenamiento preclasificados. Se requiere por lo tanto que la intuición humana no sea eliminada del todo, pues el diseñador del sistema tiene que especificar como los datos tendrán que ser representados y que mecanismos serán usados para averiguar por la caracterización de los datos. Es importante por lo tanto que se adopte en alguna medida acercamientos colaborativos entre el hombre y la maquina.

El proceso de aprendizaje para las Cadenas Oculta de Markov (HMM), consiste en aplicar técnicas de aprendizaje inductivo, en los cuales a partir de un conjunto masivos de datos recibidos durante el entrenamiento infiere una regla generalizada (patrón) que los gobierna. Esta regla es obtenida con el ajuste de parámetros estructurales del HMM ( $\lambda=(A,B,\pi)$ , más adelante se mencionará estos parámetros) constituidos fundamentalmente por valores de probabilidad que modelan el sistema real, permitiendo al diseñador tener una manera de cuantificar y permitir mejorar el desempeño de la salida del sistema de dicho modelo gracias a su constitución

numérica. El conjunto de datos de entrenamiento recibidos por el HMM por lo general son datos preclasificados que son destinados a representar una clase específica por dicho HMM. Esto manifiesta un aprendizaje supervisado de sistema, sin dejar a un lado la posibilidad de realizar una hibridación que conecte previamente a un HMM un mecanismo de aprendizaje no supervisado que pre-clasifique los datos de manera automática. Mecanismos de este tipo se mencionan clasificadores como las K-NN (Nearest Neighbor), Redes neuronales de Perceptron Multicapa, Mapas auto-organizadores de Kohonen, Growing Neural Gas, Algoritmo de K-medias en línea, entre otros.

Algunas de las técnicas de entrenamiento de HMM's para el aprendizaje, consisten en procedimientos iterativos como los desarrollados por Baum Welch (equivalente al método EM: Expectation-Maximization)<sup>8</sup> o técnicas de uso de gradientes que son utilizados para ajustes aproximados de los parámetros de probabilidad. El algoritmo de Baum Welch utiliza el conjunto de datos de entrenamiento de manera que en cada iteración actualice los valores probabilísticos del modelo  $\lambda=(A,B,\pi)$ , que mejor representen a todos los datos procesados hasta la última iteración. Cada iteración representa un avance en el aprendizaje. El resultado es la selección de los parámetros que represente un máximo de probabilidad (local o global) dentro de un espacio de soluciones (llamadas hipótesis) iterativamente. En la Figura 4 se observa este hecho.

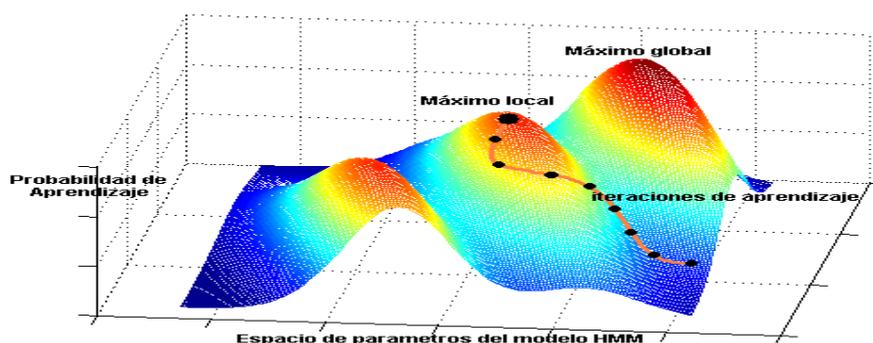


Figura 4. Espacio de parámetros para selección de probabilidades máximas<sup>9</sup>.

<sup>8</sup> Rabiner, Lawrence R. *A tutorial on Hidden Markov Models and Selected Applications in Speech Recognition*. En: *Proceeding of IEEE*. Vol. 6, No. 2. (Feb 1989).p. 264

<sup>9</sup> Imagen obtenida con: MatLab 5.3.

Más adelante en el capítulo 3.4.3 se tratará el algoritmo de Baum-Welch utilizado para el entrenamiento y ajuste de los parámetros del HMM.



**2.2.2. Acerca de la información que procesa.** Los otros aspectos importantes a la hora de elegir y diseñar un sistema inteligente, es conocer el tipo de información al cual estará sujeto el sistema y las tareas que puede ejecutar con ese tipo de información después del proceso de entrenamiento como método de aprendizaje.

Un factor importante de la información (datos de entrenamiento y prueba) es el modo en que se presenta. Los sistemas emiten señales que pueden ser presentarse de distintas formas. Una señal puede ser **continua** como las variaciones de temperatura o presión de un tanque cuyos valores están definidos para un rango de números reales; **discreta** como el alfabeto o un conjunto de colores, estados climáticos o números enteros; **estacionaria** como una fuente de potencia que emite una señal constante en forma indefinida (sus parámetros no varían con el tiempo), o incluso una imagen estática; **No estacionaria (estocástica o dinámica)** como el comportamiento del clima cuyos factores están muy dependientes del tiempo presentándose en forma aleatoria o el estado de deterioro de un equipo que presenta variaciones dinámicas temporales.

Otro aspecto importante, es la reacción de la máquina a la presencia de perturbaciones en la información, como lo es ruido, datos incompletos, ambiguos, etc. Independientemente de la forma en que se presente la información, la principal tarea de una máquina de aprendizaje en su etapa de entrenamiento, es la de obtener los patrones adecuados que caracterizan el comportamiento del proceso en base a la información recibida. Después de que este patrón sea representado por los parámetros internos de la máquina, se espera que esta acepte dentro de la información obtenida durante la ejecución de alguna tarea de reconocimiento, algunas perturbaciones como la variabilidad e incertidumbre que aparentemente podría confundir a la máquina y afectar el desempeño en la ejecución de su actividad. Un robot, por ejemplo, que tenga

la misión de seleccionar y recorrer una de varias trayectoria para transportan un objeto dentro de una planta, se le ha entrenado para discriminar entre las vías que son navegables y no navegables para un transporte seguro. Esta habilidad de discriminar es controlada por los patrones obtenidos en su etapa de entrenamiento, permitiéndole cumplir a cabalidad su tarea sin preocuparse que la información que recibe de sus sensores (cámaras de video, ultrasonido, laser, etc.), llegase a estar incompleta, errónea, tener ligeras imprecisiones, variaciones o ruido en periodos de tiempo cortos, sometiendo al robot en ese instante a tener incertidumbre de lo que observa del medio.

El tipo de información con el que interactúa las Cadena Ocultas de Markov (HMM) contiene series de datos temporales (continuos o discretos) con características aleatorias, cuyo comportamiento está regido por algún tipo de **patrón secuencial**. Esto sugiere la existencia de algún modo de dependencia o correlación en el orden de aparición de los patrones a través del tiempo<sup>10</sup> y su identificación determina el comportamiento o “personalidad” de un proceso específico. Se verá más adelante, que estos patrones secuenciales proporcionan capacidades de predicción para la clasificación de algún tipo de evento temporal.

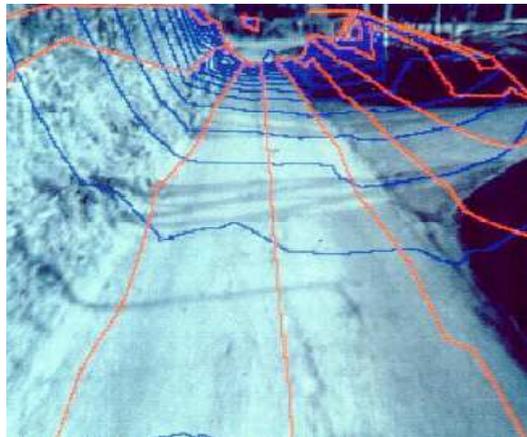
Cabe hacer una comparación entre patrones estáticos y patrones secuenciales, los cuales contienen significados diferentes. Una imagen estática puede contener características únicas que lo diferencia de otras imágenes, teniendo este un patrón estático cuyas características no cambiarán con el tiempo; Una palabra escrita por alguien, la forma de algún objeto regular o irregular, tiene un patrón estático que los diferencia de otros. Por otro lado, una secuencia de imágenes que muestran un comportamiento de los objetos dentro de ella en un orden y en función del tiempo está regido por un patrón secuencial de comportamiento temporal; La manera en que alguien escribe manualmente una secuencia de letras es única de este ente y es caracterizado por un patrón secuencial que explica como es el proceso de movimiento de la mano del escritor a través del tiempo (Handwriting recognition).

Siguiendo con el ejemplo del robot, supóngase que el HMM se comunica con el entorno a través de una cámara de video. A medida que el robot avanza, la cámara

---

<sup>10</sup> DIETTERICH, Thomas G. *Machine Learning for Sequential Data: A Review*. Oregon, USA. Oregon State University. 15P.

registra una secuencia de imágenes que cambian con el tiempo, lo que significa que desde el punto de vista del robot se detecta una secuencia de cambios en el comportamiento del entorno que se presenta en las imágenes. Como se puede observar en la Figura 5, el robot hace un análisis de las características importantes del terreno (por algún algoritmo especial de mapeo)<sup>11</sup> y el HMM evalúa la probabilidad de que el terreno analizado con las características obtenidas sea navegable o no, acorde al conocimiento adquirido en el entrenamiento. De no ser navegable, el sistema de control del robot lo hará moverse en una dirección que genere un mapa probable de ruta navegable para el HMM y continuara su camino repitiendo el proceso hasta que termine de ejecutar su tarea.



**Figura 5. Análisis de capturas seriadas de imagen del terreno por un Robot Móvil<sup>12</sup>.**

Otras aplicaciones de presentan información de similar naturaleza se mencionan por ejemplo a aquellas que se presentan conductas de comportamientos en seres humanos a partir de movimientos corporales, trayectorias de blancos o partículas, proceso de envejecimiento de equipos, estados climáticos, cambios de conducta psicológica y medica en pacientes, y muchas otras que presentes variaciones temporales con correlaciones entre información pasada y presente.

---

<sup>11</sup> En la referencia [26] se da un panorama general de las técnicas de mapeo de terrenos con robots.

<sup>12</sup> Imagen obtenida de: VILLEGAS, Juan. *Aprendizaje en maquinas (Machine Learning)*. Cursos temas selectos. Febrero 2007. p.2

En resumen, la información con el cual interactúa un HMM tiene las siguientes características:

- Puede ser continua o discreta,
- Puede contener incertidumbres por errores en sensores (datos incompletos, ambiguos) o estar afectada con ruido.
- Se manifiesta en series de datos temporales (Time-series data) regidos por un patrón secuencial.



**2.2.3 Acerca de las actividades que ejecuta.** Conociendo el tipo de información al cual tratará la máquina, se querrá saber qué clase de actividades podría ejecutar. Acorde a los ejemplos anteriores, un HMM es diseñado para realizar tareas específicas de diferente tipo.

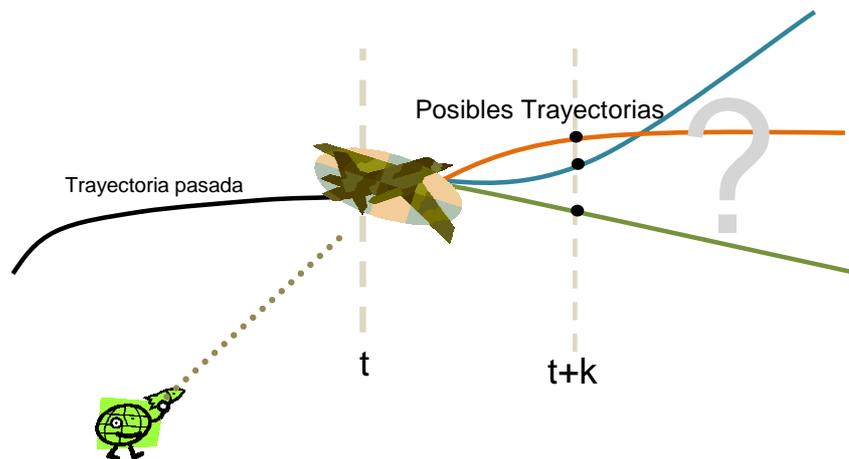
Una de las tareas que puede hacer un HMM es la **predicción** de eventos, la cual aprovecha la capacidad que tiene ese tipo de máquina para descubrir patrones secuenciales en el tiempo en forma probabilística acorde a lo ocurrido en tiempos anteriores. Las HMM's son modelos Generativos probabilísticos<sup>13</sup> para series de datos cuya finalidad es el de modelar comportamiento de sistemas reales en el **tiempo** que posteriormente serán utilizados para reproducir o generar con un margen de probabilidad dicho comportamiento. Teniendo definido los parámetros probabilísticos del modelo, se puede utilizar al HMM como un generador de datos<sup>14</sup> que presenta una simulación de los eventos que probablemente podrían ocurrir en un futuro y así tomar decisiones acertadas. Por ejemplo, En el área financiera de una compañía es muy importante conocer en forma previa (**a priori**) la ocurrencia de un evento financiero futuro que influya en sus actividades económicas utilizando un modelo matemático que

---

<sup>13</sup> ROWEIS, Sam. SCIA 2003 Tutorial: Hidden Markov Models. Toronto, June 29 2003, University of Toronto. 10P.

<sup>14</sup> Se mostrará en la sección 3.3 a un HMM como generador de observaciones.

haga una predicción a partir de los eventos pasados. En las actividades mantenimiento predictivo de una planta, es importante predecir el estado de deterioro de un equipo en tiempos posteriores de acuerdo a su historial, para programar en forma efectiva los recursos necesarios para un mantenimiento antes de que el equipo falle inesperadamente. Igualmente ocurre en el área de diagnóstico médico para enfermedades específicas; en los sistemas de control de generación de potencia eléctrica para predecir la demanda de energía en un instante específico de acuerdo a tendencias anteriores, entre otras. Por ejemplo, se muestra en la Figura 6, el problema de poder predecir la posible trayectoria de una partícula en un ambiente Tridimensional (3-D) dentro de las posibles trayectorias modeladas por varias HMM's para explicar el comportamiento de esa partícula a través del tiempo.



**Figura 6. Problema de predicción de posibles trayectorias de una nave.**

HMM también puede ejecutar tareas de **Decodificación** o descubrimiento de secuencias de eventos no visibles ni medibles. Se observará más adelante que algunos sistemas están constituidos de doble proceso estocástico, donde uno de ellos es oculto y no observable y que a la vez gobierna al otro proceso que emite variables observables. El objetivo de decodificar el comportamiento de un sistema consiste en describir una hipótesis de la secuencia o desarrollo del proceso oculto solo a través del proceso observable.

Este es el caso del reconocimiento de voz, donde a partir de una secuencia de señales acústicas (variable medible) recogidas por un transductor, se obtenga el texto de lo que fue pronunciado. Aquí la tarea de decodificación consiste en generar hipótesis de la secuencia de fonemas o silabas más probables del lenguaje (proceso no observable), a partir de estas señales acústicas pronunciadas por el sistema fonador (proceso observable).

Actividades como la **Clasificación** también son llevadas a cabo por HMM. Después de que la etapa de entrenamiento sea llevada a cabo para varias HMM's, la cual cada una representa una clase distinta de comportamiento temporal, se puede someter un nuevo evento temporal sin clasificar para ubicarlo en la categoría del HMM que sea más probable (Figura 7).

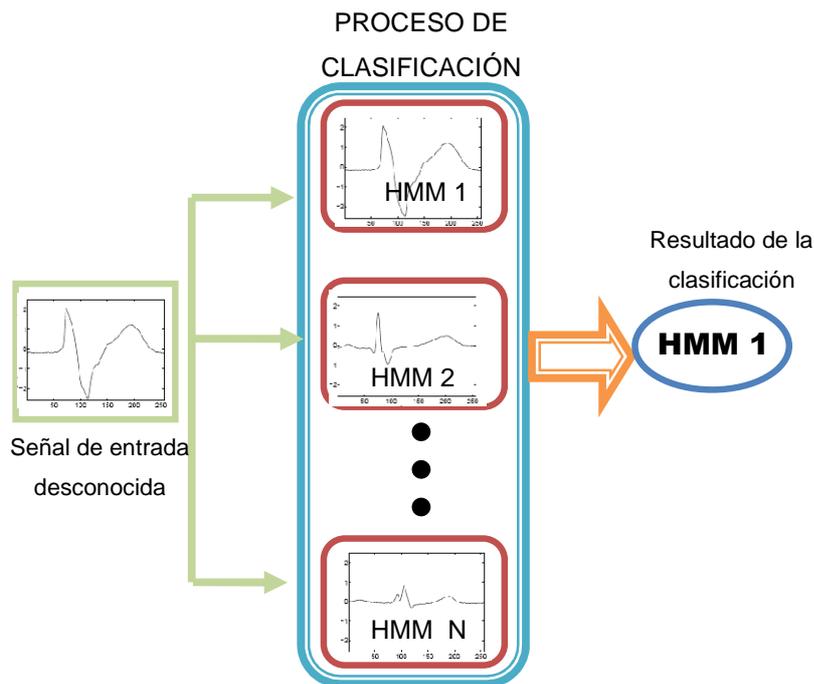


Figura 7. Clasificación de eventos temporales desconocidos con HMM.

### RESUMEN: INTRODUCCIÓN A CADENAS OCULTAS DE MARKOV

Existen problemas que requieren ser resueltos por medio de inteligencia para reducir riesgos de accidentes, tomar decisiones, hacer predicciones. etc. Se aprovechan los avances tecnológicos en los computadores para incrementar la rapidez con que se soluciona problemas. Se utilizan métodos probabilísticos para modelar procesos que cambian con el tiempo y que poseen características aleatorias.

#### 1. MECANISMO DE APRENDIZAJE DE HMM

Métodos inductivos para el ajuste de parámetros probabilísticos que representan un patrón secuencial. Uso del algoritmo de Baum-Welch y procedimiento EM (Expectation-maximization)

#### 2. INFORMACIÓN QUE MANEJA HMM:

- Puede ser continua o discreta,
- Puede tolerar incertidumbres por errores en sensores (datos incompletos, ambiguos) o estar afectada con ruido.
- Se manifiesta en series de datos temporales (Time-series data).

#### 3. TAREAS DE UN HMM:

- Predicción: ¿Que es lo más probable que pueda suceder después?
- Decodificación: ¿Cuál es la secuencia de eventos no observables más probable?.
- Generación de datos a través de su modelo (Simulación)
- Reconocimiento y Clasificación ¿Si es compatible la información nueva con el modelo?

### 3. CADENAS OCULTAS DE MARKOV: ESTRUCTURA Y ANÁLISIS

Conociendo las potencialidades de la teoría de la probabilidad para modelar procesos estocásticos donde predomina la incertidumbre, se mostrará de manera general la estructura de las HMM's para entender específicamente las intenciones de esta máquina y corroborar su mecanismo de aprendizaje, la naturaleza de la información de entrada y las tareas que puede ejecutar.

Para poder entender la estructura de las HMM's, es pertinente hacer una introducción de las Cadenas de Markov (no ocultas), en donde esta última presenta propiedades que serán utilizadas en las HMM's.

#### 3.1. CADENAS DE MARKOV<sup>15</sup>

Una cadena de Markov es un proceso estocástico que modela el comportamiento de un proceso con características aleatorias en el transcurso del tiempo, utilizando una familia de variables aleatorias parametrizadas. Un proceso aleatorio se caracteriza por presentar eventos que no siguen una regla determinada y por lo tanto no es fácil definir con certeza una secuencia de eventos previamente (a priori) en forma segura antes de que estos ocurran. En contraste, se puede definir con certeza por ejemplo, la secuencia de eventos de operación de un semáforo (Verde–Amarillo-Rojo-Verde) en forma previa antes que termine el ciclo completo. En la Figura 8 se observa este proceso representado como una máquina de estados.

---

<sup>15</sup> Tomado de: Rabiner, Lawrence R. A tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. En: Proceeding of IEEE. Vol. 6, No. 2. (Feb 1989),p.257-259.

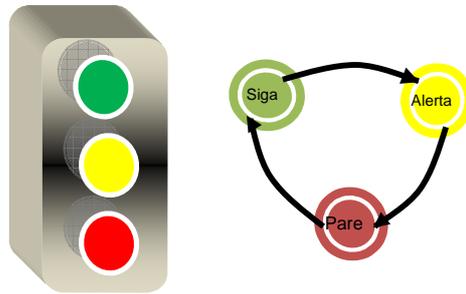
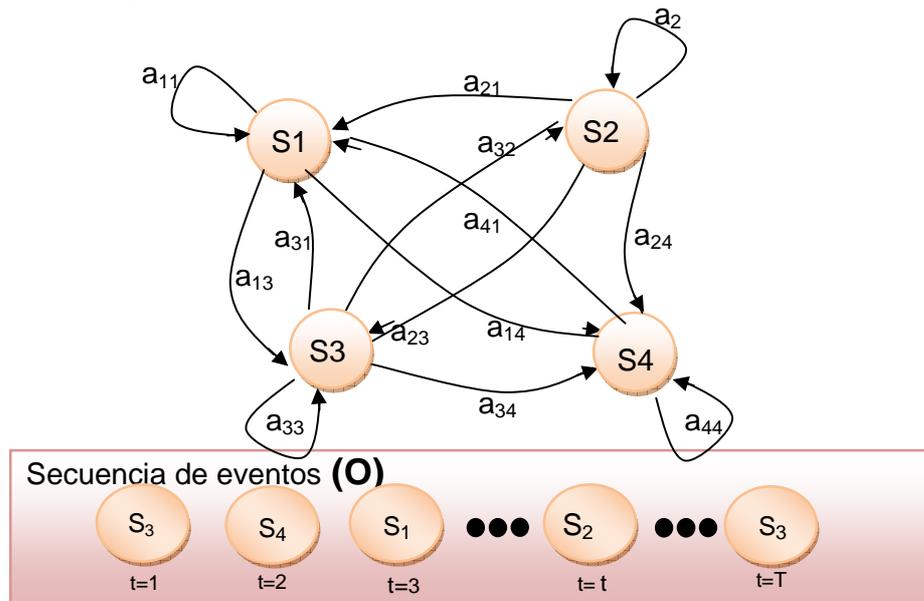


Figura 8. Proceso determinista: Iluminación del Semáforo.

Este sistema trabaja bajo reglas deterministas que le indican la secuencia de eventos exacta a seguir para que la operación tenga sentido. En las Cadenas de Markov ocurre lo contrario: se conoce cuales son los eventos posibles de un proceso pero no se conoce con absoluta certeza como se presentaran a través del tiempo.

Considere el siguiente modelo de Markov en donde se representa un proceso aleatorio en el tiempo por medio de un diagrama de transición de eventos. Se puede observar en la Figura 9 que si en un instante dado ocurre un evento, en el próximo instante puede ocurrir cualquiera de los otros eventos de acuerdo al valor de probabilidad de las conexiones  $a_{ij}$ .

Figura 9. Maquina de estados Cadena de Markov.



En el modelo se consideran 4 eventos que han sido definidos a partir de algún sistema real. A estos eventos se les conoce como **Estados** del proceso aleatorio los cuales están conectados entre sí por medio de unos valores de probabilidad de transición.

En cualquier instante de tiempo puede estar en cualquiera de los  $N$  estados del sistema  $S_1, S_2, S_3 \dots S_N$ , (En este caso  $N=4$ ), y en intervalos de tiempo discretos (segundos, minutos, días, años, etc) regularmente espaciados. El sistema puede presentar cambios de estado (incluso volver a sí mismo), de acuerdo a un conjunto de probabilidades relacionadas al estado previo. Los instantes de tiempo asociados a los cambios de estado se denotan  $t=1,2,3\dots$ , y por lo tanto la denominación de un estado en un instante  $t$  cualquiera es  $q_t$ , que según la **Figura 9** por ejemplo  $q_2= S_4$  o  $q_1= S_3$ . Cuando se presenta una secuencia de eventos como se observa en la **Figura 9**, este conjunto se denota  $O = \{q_1, q_2, q_3, \dots q_t, q_{t+1}, \dots q_{T-1}, q_T\}$ .

El hecho de que un estado sea seleccionado en cierto instante  $t$ , se infiere que sus estados anteriores contribuyeron a que fuera así. La expresión:

$$P[q_t = S_j | q_{t-1} = S_i, q_{t-2} = S_k, \dots]$$

Define la probabilidad (condicional) de que en el instante  $t$  sea seleccionado el estado  $S_j$ , con la condición de que en instantes anteriores  $t-1, t-2, \dots$ , sean seleccionados los estados  $S_i, S_k, \dots$ , respectivamente. Una cadena de Markov de  $n$  orden tiene en cuenta los  $n$  estados seleccionados anteriormente para seleccionar un nuevo estado. Entre más alto sea el orden de la Cadena mas “memoria” tendrá del pasado para seleccionar un estado en el presente.

Para el caso especial de una Cadena de Markov de primer orden, la expresión anterior solo se reduce a tener en cuenta el estado en el instante anterior  $t-1$  de tal manera que  $P[q_t = S_j | q_{t-1} = S_i]$ . Esta expresión obedece a la propiedad de Markov que dice “Conocido el estado del proceso en un momento dado, su comportamiento futuro no depende del pasado”. Dicho de otro modo, “dado el presente, el futuro es independiente del pasado”.

Esta propiedad esta cuantificada por medio de los valores probabilidad de transición entre 2 estados  $i, j$  como se observa en la figura **Figura 9**, representados por:

## Hidden Markov Models (HMM's) y aplicaciones

---

$$a_{ij} = P[q_t = S_j | q_{t-1} = S_i] = P[S_j | S_i] \text{ donde } 1 \leq i, j \leq N$$

Los cuales deber cumplir las siguientes restricciones:

- Las transiciones entre un par de estados  $i, j$  por ser valores de probabilidad deben ser positivas y menores que la unidad. ( $0 \leq a_{ij} \leq 1$ ).
- La suma de todas las transiciones que salen de un estado  $S_i$  hacia otros estados deber ser igual a la unidad  $\sum_{j=1}^N a_{ij} = 1$ .

Otra propiedad que se asume en las cadenas de Markov es que los valores de probabilidad de transición son independientes en el tiempo (estacionarios). La expresión:

$$a_{ij} = P[q_t = S_j | q_{t-1} = S_i] = P[q_{t+l} = S_j | q_{t+l-1} = S_i]$$

indica que la probabilidad de transitar de un estado  $S_i$  a otro estado  $S_j$  se mantiene constante sin importar cuando se ejecute dicho cambio.

Las transiciones entre estados que se muestran en la figura por medio de flechas se pueden representar por medio de una matriz  $\mathbf{A}$  como la siguiente:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \text{ Suma igual a 1}$$

Donde las filas representan el estado  $i$  (desde) en un instante  $t$ , y las columnas el estado  $j$  (hacia) en un estado  $t+1$ .

Según el diagrama de la Figura 9 se muestra además que no existen transiciones  $a_{12}$ ,  $a_{42}$ , y  $a_{43}$  lo cual se entiende que no existe ninguna probabilidad (cero) de que ocurran estas transiciones.

Con fines ilustrativos, supóngase un sistema conformado por un motor de velocidad variable que realiza algún tipo de función en un proceso y se desea modelar el comportamiento de la velocidad a través del tiempo. Para casos prácticos se identifica tres estados básicos (espacio de estados) de la variable velocidad entre ellos:

**Estado 1 ( $S_1$ ):** Velocidad en Aumento (VA); **Estado 2 ( $S_2$ ):** Velocidad en Descenso (VD); **Estado 3 ( $S_3$ ):** Velocidad Constante (VC).

En este caso no será de interés la magnitud de la velocidad, sino algunos de los 3 estados anteriores que ocurren en el tiempo, entendiendo a estos cambios como algún fenómeno inherente aleatorio dentro del sistema. Se supone además que estos estados se pueden observar fácilmente por cualquier medio.

Supóngase que se desea analizar los eventos en intervalos de tiempo de 1 minuto, es decir que en cada instante de tiempo se registra el estado del motor. En la Figura 10 se muestra una grafica de una secuencia posible de estados de la velocidad del motor en el tiempo. Cabe aclarar que se podrían presentar otra secuencia de estados diferentes a esta (con pequeñas variaciones), pero siguiendo el mismo patrón de comportamiento temporal acorde a sus parámetros de probabilidad (invariabilidad en el tiempo) que obedecen al funcionamiento normal del motor dentro del sistema. Junto a la gráfica se observa el diagrama de transición de estados de este sistema, cuyos valores de probabilidad de transición se suponen que ya han sido definidos previamente en el entrenamiento de acuerdo a comportamientos pasados del sistema.

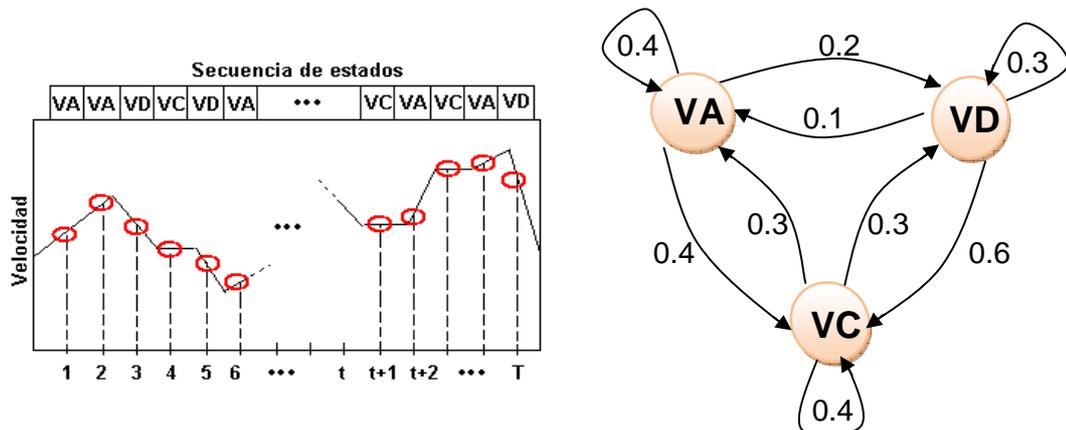


Figura 10. Modelo Cadena de Markov para motor de velocidad variable

La matriz de probabilidad de este sistema es el siguiente:

$$A = \begin{bmatrix} 0.4 & 0.2 & 0.4 \\ 0.1 & 0.3 & 0.6 \\ 0.3 & 0.3 & 0.4 \end{bmatrix}$$

Teniendo una matriz como la anterior que modela el posible comportamiento del motor en el tiempo se podrían resolver las siguientes cuestiones:

- 1) Si se supone que el estado de la velocidad en el primer instante ( $t=1$ ) está en incremento (**VA: estado 1**), *¿Cuál sería la probabilidad de acuerdo al modelo (matriz A) que la velocidad en los siguientes 9 instantes sea " VC, VA, VA, VD, VC, VD, VC, VD, VA" ?*

Se define formalmente la secuencia de observaciones incluyendo el primer estado **VA** así:

$$O = \{q_1, q_2, q_3, q_4, q_5, q_6, q_7, q_8, q_9, q_{10}\} = \{S_1, S_3, S_1, S_1, S_2, S_3, S_2, S_3, S_2, S_1\}$$

Téngase en cuenta que solo se necesita el estado actual **VA** para predecir que tan probable es que ocurran los 9 eventos siguientes (Propiedad de Markov de primer orden).

La probabilidad de que ocurra **O** acorde al modelo se expresa:

$$\begin{aligned} P(O|Modelo) &= P[S_1, S_3, S_1, S_1, S_2, S_3, S_2, S_3, S_2, S_1|Modelo] \\ &= P[S_1] * P[S_3|S_1] * P[S_1|S_3] * P[S_1|S_1] * P[S_2|S_1] * P[S_3|S_2] * P[S_2|S_3] * P[S_3|S_2] * P[S_2|S_3] \\ &\quad * P[S_1|S_2] \\ &= \pi_1 * a_{13} * a_{31} * a_{11} * a_{12} * a_{23} * a_{32} * a_{23} * a_{32} * a_{21} \\ &= 1 * 0.4 * 0.3 * 0.4 * 0.2 * 0.6 * 0.3 * 0.6 * 0.3 * 0.1 = 3.1104 * 10^{-5} \end{aligned}$$

Donde  $\pi_1$  Indica la probabilidad de que el estado inicial sea **S<sub>1</sub>**, es decir, en general  $\pi_i = P[q_1 = S_i]$ .

- 2) Si se supone que el modelo se encuentre en cierto estado de velocidad, *¿cuál es la probabilidad de que permanezca en ese estado por exactamente **d** minutos?*

Esta situación se puede analizar a partir de la secuencia de observaciones:

$$O = \left\{ \underset{1}{S_1}, \underset{2}{S_2}, \underset{3}{S_3}, \dots, \underset{d}{S_d}, \underset{d+1}{S_{d+1}} = S_i \right\}$$

Cuya probabilidad se puede calcular así:

$$P(O|\text{Modelo}, q_1 = S_i) = (a_{ii})^{d-1}(1 - a_{ii}) = p_i(d)$$

La expresión  $p_i(d)$  es la función de densidad de probabilidad de duración  $d$  en estado  $i$ .

- 3) Finalmente sería interesante calcular cuantos minutos se puede esperar que las observaciones permanezcan en un mismo estado de velocidad. La respuesta a esta pregunta está dada por:

$$\bar{d}_i = \sum_{d=1}^{\infty} d p_i(d) = \sum_{d=1}^{\infty} d (a_{ii})^{d-1}(1 - a_{ii}) = \frac{1}{1 - a_{ii}}$$

Por lo tanto se podría esperar que el motor se encuentre incrementando su velocidad (**VA**) por 1.67 minutos; se encuentre reduciendo su velocidad (**VD**) por 1.43 minutos; o se encuentre en velocidad constante (**VC**) por 1.67 minutos.

Como el ejemplo anterior, existe varios que explican la intenciones de las cadenas de Markov, entre estos, el lanzamiento de monedas con dos estados posibles (cara y sello); el lanzamiento de dados con 6 estados posibles; comportamientos climáticos con 3 estados (soleado, nublado o lluvioso); estado de deterioro de un artefacto (Nuevo, funcional, defectuoso, no funcional), etc.

Se observa la utilidad de las cadenas de Markov para modelar procesos que a pesar de ser aleatorios en el tiempo se pueden parametrizar y conocer su comportamiento secuencial con un margen de probabilidad previa. Sin embargo muchas de las aplicaciones que interesan a los científicos requieren de métodos más potentes debido a que la naturaleza de algunos procesos contiene estados que no son visibles, ni medibles, ni mucho menos accesibles, inhabilitando la posibilidad de utilizar Cadenas de Markov. Un método alternativo, es el de seleccionar variables accesibles y medibles que estén probabilísticamente relacionados a los estados ocultos del sistema. Este método es característico de las HMM's, los cuales se presentan a continuación.

### 3.2. CADENAS OCULTAS DE MARKOV (HMM)<sup>16</sup>

Hasta este punto se conoce las Cadenas de Markov, en el cual cada estado corresponde a una observación del conjunto  $O$ . Este tipo de modelo restringe frecuentemente la aplicabilidad a diferentes problemas de interés del mundo real. La mayoría de procesos a modelar, experimentan cambios de estado los cuales sugieren algún significado (simbólico o no simbólico) relevante e indispensable a los científicos que se dedican a analizar los impactos que representan estos cambios. Sin embargo, para muchas aplicaciones, estos estados no son fácilmente identificables a primera vista (no observable), lo que conlleva a la inspección de variables externas observables (cuantitativas o cualitativas) relacionadas a los estados ocultos del proceso y a partir de estas inferir la secuencia de estados por los cuales se ha desarrollado. Es por esto que un conjunto de variables observables del proceso con valores o características definidas en un instante dado, puede dar información del posible estado en que se encuentra dicho proceso, teniendo en cuenta además que las observaciones pueden ser ruidosas, ser imprecisas, erróneas o en el peor de los casos estar incompletas.

HMM surgió de la necesidad de estudiar el problema de caracterizar procesos estocásticos en el cual las observaciones obtenidas estaban incompletas. La única forma de no prescindir de las observaciones como partes fundamentales de los sistemas de predicción, era adoptar un nuevo pensamiento en donde se fundamentaba en que las observaciones seguían un patrón de comportamiento oculto y por tanto la ausencia de algunas observaciones ya no era tan relevante. Este patrón se desarrolla en la secuencia de estados ocultos en los cuales dependiendo de las características del proceso a modelar, tengan un comportamiento o secuencia probablemente definida. De aquí surgió la idea de HMM como un sistema doblemente estocástico, es decir, que presentaba aleatoriedad tanto en las observaciones, como en la secuencia de estados ocultos que gobernaban dichas observaciones. Lo más interesante aun es que los dos procesos estocásticos deben ser definidos en la etapa de entrenamiento solamente a través de las observaciones. Como se mencionó, su estructura consta de

---

<sup>16</sup> Ibid.,p.260-261.

dos capas estocásticas. La primera capa consiste en un proceso de estados ocultos secuenciales en el cual no son directamente observables y que obedecen a las reglas de transición de los Modelos de Markov que se mencionaron anteriormente. La segunda capa es un proceso de emisión de símbolos de observación, en el cual es físicamente observable y tienen cierta función de distribución de probabilidad dependiendo del estado oculto en el que se encuentra actualmente.

A continuación se mencionan los componentes de un HMM.

**3.2.1. Componentes de un HMM.** Es de vital importancia conocer los elementos de un HMM, de manera que se tenga la capacidad de identificar cada uno de estos en una aplicación específica que se quiera desarrollar. Las Cadenas Ocultas de Markov están formadas por 5 elementos fundamentales, como se observa en la Figura 11 y se mencionan:

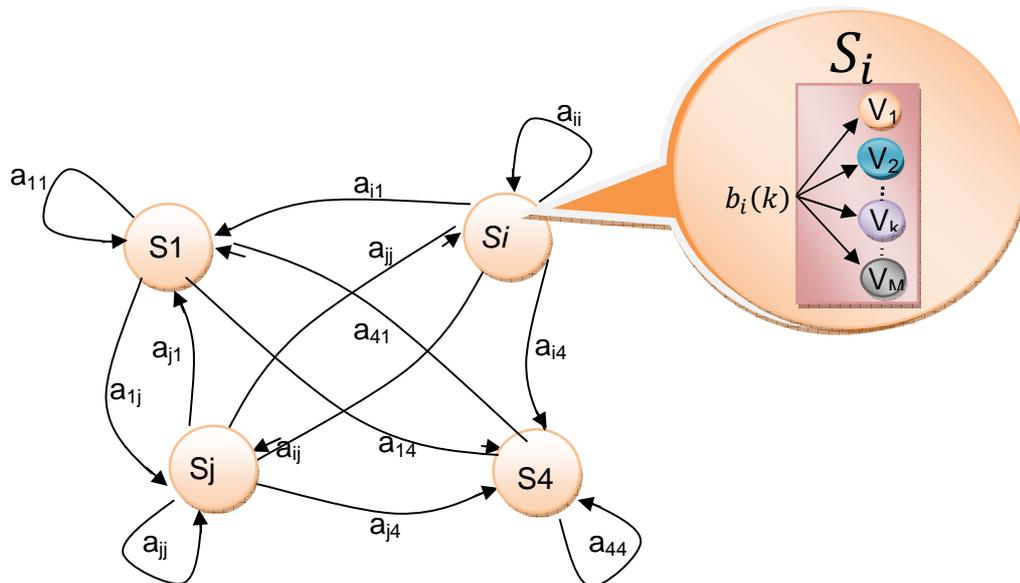


Figura 11. Cadena Oculta de Markov. Estados y distribución de observaciones

**a) Número de estados (N):** Aunque sean estados ocultos, para muchas aplicaciones prácticas hay un significado relevante adherido al estado o al conjunto de estados del modelo determinado por el usuario y las características naturales de la

aplicación. Esto quiere decir, que cada estado, dependiendo de la aplicación, puede tener una etiqueta identificable y significativa para el usuario dentro del proceso. En otros casos no se puede etiquetar y no tener ningún significado. La definición de su número para una aplicación específica es absolutamente dependiente de su propia naturaleza, además del conocimiento previo del diseñador con respecto a aquella aplicación.

El conjunto de estados de la maquina se representan:

$$S = \{S_1, S_2, S_3 \dots S_N\}$$

La selección de un estado  $S_i$  en un instante  $t$  quedara denotado por:

$$q_t = S_i \quad \text{donde} \quad 1 \leq i \leq N$$

**b) Número de distintas observaciones posibles por estado (M).** Las observaciones son las salidas físicas de un proceso o sistema que ha sido modelado. El conjunto de símbolos que generan las observaciones se denota:

$$V = \{V_1, V_2, V_3 \dots V_k \dots V_M\}$$

En este vector se define el dominio de todos los M posibles símbolos que podrían tomar las observaciones que se presentan en el proceso a modelar.

Cuando del proceso se obtienen una secuencia de observaciones de longitud  $T$ , estas se definen:

$$O = O_1 O_2 O_3 \dots O_t \dots O_{T-1} O_T$$

En la cual cada elemento  $O_t$  del vector  $O$  es equivalente a un símbolo  $V_k$  del Vector  $V$  en un instante  $t$ . Naturalmente, dentro del conjunto de Observaciones  $O$  se pueden repetir elementos  $V_k$  sin ninguna restricción.

**c) Matriz de Distribución de probabilidad de transiciones entre estados (A).** Sus componentes son valores de distribuciones de probabilidad para cambiar de un estado oculto a otro igual como se observó en una Cadena de Markov. Matemáticamente se expresa:  $A = \{a_{ij}\}$  Una matriz con dimensión  $N \times N$

y sus elementos:  $a_{ij} = P[q_{t+1} = S_j | q_t = S_i] \quad 1 \leq i, j \leq N$

Entiéndase  $a_{ij}$ , como un valor de la matriz  $A$  en el cual cuantifica la probabilidad de pasar del estado  $S_i$  en un tiempo  $t$ , al estado  $S_j$  en un tiempo posterior  $t+1$ .

Dependiendo de la topología de la HMM, existe la posibilidad que no existan conexiones entre estados  $i$  y  $j$  ( $a_{ij} = 0$ ), aunque es indispensable que un estado tenga al menos una conexión con otro estado ( $\sum_{j=1}^N a_{ij} \neq 0$ )

**d) Matriz de Distribución de probabilidad para la selección de los símbolos de observación ( $B$ )**, definidos por una distribución de probabilidad cuando se está en un estado determinado para la selección de observaciones. Matemáticamente se expresa:  $B = \{b_j(k)\}$  Una matriz de  $N \times M$

y sus elementos:  $b_j(k) = P[v_k \text{ en } t | q_t = S_j] \quad 1 \leq j \leq N, 1 \leq k \leq M$

Esto es, la probabilidad que el símbolo  **$k$ -ésimo** de todas las  **$M$**  posibles del vector  **$V$**  sea seleccionado en un tiempo  **$t$**  cuando en ese instante se encuentre en el estado  **$S_j$** .

Las restricciones para esta matriz son:

- Los valores de distribución deben ser positivos y menores que 1:

$$b_j(k) \geq 0$$

- Dentro de un estado  $S_j$ , la distribución de probabilidad para la selección de símbolos debe estar distribuidos entre ellos de manera tal que la suma de cada uno de estas probabilidades sea igual a la unidad:

$$\sum_{k=1}^M b_j(k) = 1$$

**e) Matriz Distribución probabilística de estado inicial ( $\pi_i$ )**. Una matriz con dimensiones  $N \times 1$  definido así:

$$\pi = \{\pi_i\} \quad \text{donde} \quad \pi_i = P[q_1 = S_i] \quad 1 \leq i \leq N$$

Son valores que determinan la probabilidad de que el primer estado oculto de inicio  $q_1$  sea el estado  $S_i$  en  $t=1$ .

### 3.3. HMM COMO GENERADOR DE OBSERVACIONES<sup>17</sup>

Un HMM en el cual se ha sido definido después del entrenamiento con todos los 5 componentes, lo capacita para que genere una secuencia de  $T$  observaciones  $O = \{O_1, O_2, O_3 \dots O_T\}$  (donde cada observación  $O_t$  es uno de los símbolos del vector  $V$ ) que puede ser bastante representativas del sistema real la cual esta modelando. Esto es equivalente a hacer una simulación para conocer en forma previa el comportamiento del sistema con un margen de probabilidad.

El ciclo de funcionamiento de un HMM para generar una secuencia  $O$  se puede entender por medio de los siguientes pasos (Figura 12):

- 1) De acuerdo a la distribución de estado inicial  $\pi$  se elige  $q_1 = S_i$ , donde  $i$  puede ser cualquier estado oculto de 1 a  $N$ .
- 2) En el instante que se selecciona el primer estado  $S_i$  de acuerdo a  $\pi$ , se asigna a esta selección un instante  $t = 1$ .
- 3) Estando en el estado seleccionado, se obtiene la observación  $O_1 = v_k$  de acuerdo a las distribución de probabilidad de símbolos  $b_i(k)$ .
- 4) Luego de haberse obtenido la observación correspondiente a  $t = 1$ , se transita a un nuevo estado en  $q_{t+1} = S_j$ , de acuerdo a la distribución de probabilidad de transiciones entre estados estando aun en el estado  $S_i$  es decir  $a_{ij}$ .
- 5) En el nuevo estado, se asigna  $t = t + 1$ , y se repite el procedimiento de selección de observaciones y posteriormente al tránsito de un nuevo estado. Estos pasos continúan si  $t < T$ , de lo contrario finaliza el proceso de generación de observaciones teniéndose la totalidad de  $T$  observaciones.

---

<sup>17</sup> Ibid.,p.261.

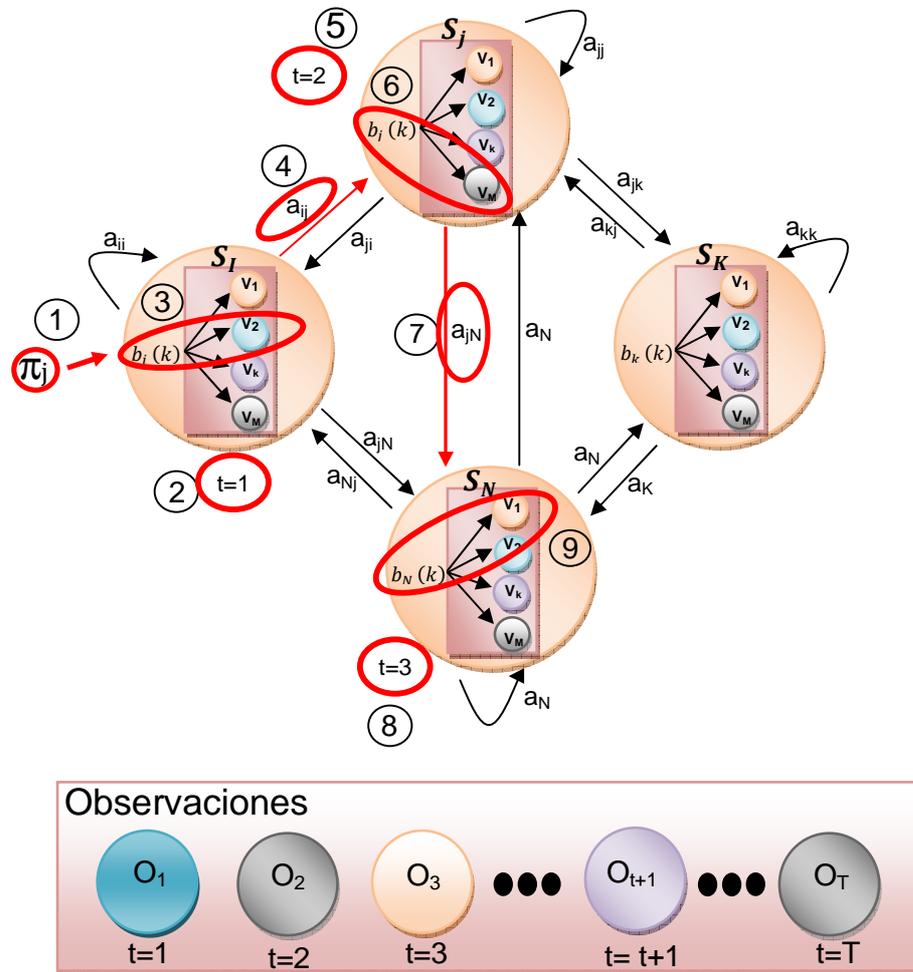


Figura 12. HMM como generador de observaciones

Como puede observarse, la especificación completa de un HMM requiere de la definición de parámetros estructurales del modelo, es decir, el número de estados  $N$ , y el conjunto de símbolos que servirán para la representación de observaciones; y un conjunto medidas probabilísticas  $A, B, \pi$ .

Para casos prácticos, estas matrices que contiene parámetros de medida probabilística se reúnen en un solo grupo, debido a que estos frecuentemente se requieren de manipulación de sus valores para realizar aproximaciones a un modelo de un sistema real.

El grupo de parámetros de medida probabilística se definirán como el **modelo de HMM** así:

$$\lambda = (A, B, \pi)$$

### 3.4. LOS TRES PROBLEMAS BÁSICOS EN HMM's Y SUS SOLUCIONES<sup>18</sup>

Dado las especificaciones de modelo  $\lambda$  de un HMM como se vio en la sección anterior, se encuentra ésta dispuesta para enfrentar los 3 problemas básicos de interés permitiéndole realizar distintas tareas aplicables del mundo real.

Es posible que una aplicación en particular requiera de la solución de más de uno de estos problemas dependiendo del uso que se le dará al modelo.

Los tres problemas básicos son los siguientes:

**3.4.1. Problema 1: Evaluación.** Dado una secuencia de observaciones  $O = O_1, O_2, \dots, O_T$  (adquiridas del mundo real o simuladas), y dado un modelo definido de HMM con sus parámetros de medida probabilística  $\lambda = (A, B, \pi)$ , ¿cómo se puede calcular de forma eficiente  $P(O|\lambda)$ ?, entiéndase esta como la probabilidad de que la secuencia de observaciones  $O$ , pueda ser asimilada o generada por el modelo  $\lambda$  del HMM definido.

---

<sup>18</sup> Ibid., p.261-265.

A este problema se le denomina *Problema de evaluación*, debido a que calcula que tan probable es que una secuencia de observaciones  $O$  dado, pueda ser representado por el HMM analizado.

Para la solución de este problema, en teoría, la probabilidad puede ser calculada directamente así:

$$P(O|\lambda) = \sum_{\text{todo } q} P(O|Q, \lambda)P(Q|\lambda)$$

$$Y: P(O|Q, \lambda)P(Q|\lambda) = \{b_{q_1}(O_1)b_{q_2}(O_2)b_{q_3}(O_3) \dots b_{q_T}(O_T)\} \{\pi_{q_1}a_{q_1q_2}a_{q_2q_3} \dots a_{q_{T-1}q_T}\}$$

Donde el termino  $P(O|Q, \lambda)P(Q|\lambda)$  se entiende como la probabilidad que las Observaciones dadas  $O$ , puedan ser representados por una secuencia determinada arbitrariamente de estados ocultos  $q_1, q_2, q_3 \dots q_T$ , multiplicado por la probabilidad de que ocurran estos estados ocultos acordes al modelo definido  $\lambda$ . La definición de este término implica determinar en forma arbitraria una secuencia de estados ocultos para el cálculo de esta probabilidad, cuando en realidad la secuencia de estados ocultos nunca se pueden “observar” en un proceso y por lo tanto son desconocidas. Se debe entonces, calcular este término tantas veces como posibles combinaciones de estados ocultos se puedan obtener dependiendo del número de observaciones a evaluar.

El número de multiplicaciones para este cálculo requiere de aproximadamente  $2T$ , donde  $T$  es la longitud de la secuencia de observaciones  $O$ . Luego como se observa en la expresión anterior, los resultados de computación del término  $P(O|q, \lambda)P(q|\lambda)$  para todas las posibles combinaciones de estados ocultos deben ser sumados ( $\sum_{\text{todo } q} P(O|Q, \lambda)P(Q|\lambda)$ ). Es decir la suma es repetida  $N^T$  veces, donde  $N$  es el numero de estados ocultos. Como ejemplo, supóngase un número relativamente pequeño de  $N$  estados y  $T$  observaciones, tal como  $N=3$  y  $T= 100$ . El número de operaciones requeridas es de  $2T \times N^T = 2 \times 100 \times 3^{100} = 10^{49}$ . Este número de operaciones es computacionalmente imposible. A causa de esto se requiere de un algoritmo mucho más eficiente para la solución de este problema de evaluación.

Existen 2 algoritmos para la solución de este problema: **Algoritmo Forward** y **Algoritmo Backward**.

**3.4.1.1. Algoritmo Forward.** Utiliza la inducción para reducir en forma significativa el proceso de computación poco factible que se observó anteriormente. Nótese que al tener  $N$  estados ocultos de los cuales son utilizados para construir una secuencia de estados oculta arbitraria  $q^{arb}$  para el cálculo de probabilidad, se pueden presentar otras posibles combinaciones  $q^s$  que no difieren mucho de  $q^{arb}$  y que posiblemente presente valores equivalentes de probabilidad. Entonces, ¿Para qué repetir las mismas computaciones una y otra vez cuando se hacen ligeras variaciones en la combinación  $q^{arb}$ ? Para tomar ventaja de de estas repeticiones se define la variable **Forward**;

$$\alpha_t(i) = P(o_1 o_2 o_3 \dots o_t, q_t = Si | \lambda)$$

Esto es, la probabilidad de que la secuencia de observaciones **parciales hasta el tiempo  $t$** ,  $\{O_1, O_2 \dots O_t\}$ , ( $1 \leq t \leq T$ ) se presente cuando se encuentra en el estado oculto  $i$ , dado el modelo  $\lambda$ . Si los valores de  $\alpha_t(i)$  para todos los estados  $Si$  ( $1 \leq i \leq N$ ) son conocidos en un tiempo dado  $t$ , luego los valores de  $\alpha_{t+1}(j)$  para todos  $j$  ( $1 \leq j \leq N$ ) puede ser computado por inducción usando los valores de  $\alpha_t(i)$ ,  $a_{ij}$  y  $b_j(O_{t+1})$ . Cada paso de inducción desde  $t$  hasta  $t + 1$  requiere de  $N \times N$  operaciones, por lo tanto, el número total de operaciones que requiere es de  $N^2 T$  – Mucho más pequeño que  $2T \times N^T$  del caso infactible anterior.

Los pasos para el desarrollo de este algoritmo se resume en 3:

1) Iniciación:

$$\alpha_1(i) = \pi_i b_i(O_1) \quad \text{donde} \quad 1 \leq i \leq N$$

El cual corresponde a todos los valores de  $\alpha$  de cada uno de los estados para la primera observación.

2) Inducción:

$$\alpha_{t+1}(j) = \left[ \sum_{i=1}^N \alpha_t(i) a_{ij} \right] b_j(O_{t+1}) \quad (1 \leq t \leq T-1) \quad \text{y} \quad (1 \leq j \leq N)$$

El paso de iniciación de la variable forward indica la probabilidad del estar en el estado  $S_i$  y obtener la observación inicial  $O_1$ . El paso de inducción, el cual es el corazón del los cálculos de la variable Forward se muestra en la **¡Error! No se encuentra el origen**

**de la referencia.** La figura muestra como un estado  $S_j$  puede ser escogido en un tiempo  $t+1$  de los  $N$  posibles estados,  $S_i$  en un tiempo anterior  $t$ . Desde que  $\alpha_t(i)$  sea la probabilidad del evento conjunto que  $O_1 O_1 \dots O_t$  sea observada, y que el estado en un tiempo  $t$  sea  $S_i$ , el producto  $\alpha_t(i)a_{ij}$  es entonces, la probabilidad de el evento conjunto que  $O_1 O_1 \dots O_t$  sea observada, y el estado  $S_j$  sea buscado en un tiempo  $t+1$  a través de  $S_i$  en un tiempo  $t$ . Sumando todo este producto a través de todos los  $N$  posibles estados  $S_i$  en un tiempo  $t$  resulta en la probabilidad de  $S_j$  en un tiempo  $t+1$  con todos los anteriores observaciones parciales previas. Una vez que esto ocurre y el estado  $S_j$  sea conocido, es fácil ver que  $\alpha_{t+1}(j)$  es obtenido teniendo en cuenta la observación  $O_{(t+1)}$  en el estado  $j$ , multiplicando la cantidad sumada por la probabilidad  $b_j(O_{t+1})$ . La computación de la ecuación de inducción se lleva a cabo para todos los estados  $j$  ( $1 \leq j \leq N$ ), para un instante  $t$ ; la computación es luego iterada para  $t= 1,2,\dots T-1$ .

### 3) Finalización:

Finalmente, la ecuación de finalización da el cálculo al cual estábamos esperando como la suma de todas las variables forward  $\alpha_T(i)$ . Este es el caso:

$$P(O|\lambda) = \sum_{i=1}^N \alpha_T(i)$$

Para entenderse el funcionamiento de este y otros algoritmos que se presentaran más adelante, se hará uso gráfico de una estructura llamada **trellis**, (lattice) en el cual muestra todos los estados ocultos del sistema en columnas para cada uno de los instantes  $t$  y sus conexiones a otros estado ocultos (transiciones) en instantes superiores  $t+1$  hasta  $T$ , para facilitar el análisis con el arroj de observaciones para cada instante sin importar su longitud.

En un instante  $t=1$  la primera columna del trellis, se calcula los valores de  $\alpha_1(i)$ , ( $1 \leq i \leq N$ ). En tiempos superiores  $t = 2,3,\dots,T$ , solo se necesita calcular los valores de  $\alpha_t(j)$ , ( $1 \leq j \leq N$ ), donde cada calculo envuelve solo  $N$  valores previos de  $\alpha_{t-1}(i)$  porque cada uno de los  $N$  estados es seleccionado por los mismos  $N$  estados de la columna de tiempo anterior.

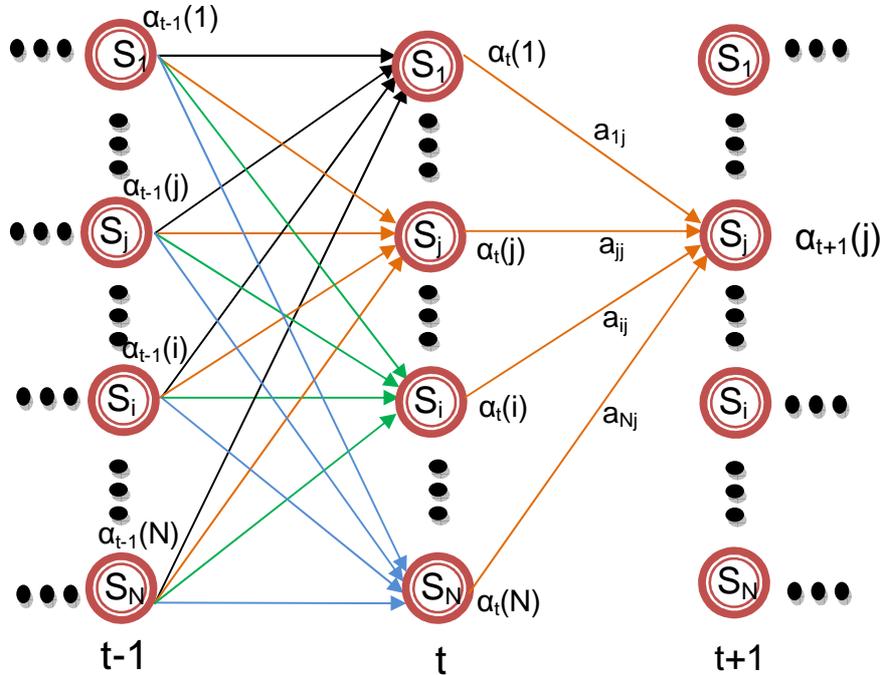


Figura 13. Ilustración del Trellis del proceso de cálculo de la variable Forward.

**3.4.1.2. Algoritmo Backward.** Este algoritmo tiene la misma filosofía que el algoritmo forward en cuanto al mejoramiento de la eficiencia de cómputo de los valores de probabilidad. Aunque este algoritmo también soluciona el problema 1, no es absolutamente necesario para la solución de este problema. Sin embargo, se hace mención de este algoritmo debido a será útil para la solución del problema 3 (Entrenamiento) que se presentará más adelante.

La variable Backward se define como:

$$\beta_t(i) = P(O_{t+1} O_{t+2} \dots O_T | q_t = S_i, \lambda)$$

Donde  $\beta_t(i)$  expresa cual es la probabilidad de que la secuencia de observaciones parciales desde un instante superior  $t+1$  hasta la última observación  $T$ , cuando se

llega a el estado  $S_i$  en un instante  $t$ , de acuerdo al modelo  $\lambda$ . Esta vez, los valores de  $\beta_t(i)$  a diferencia de  $\alpha_t(i)$  son inducido en forma reversa, analizando las observaciones desde la última, en el instante  $t=T$ , hasta la primera, en  $t=1$ . El número de computaciones requeridas es, igual que el algoritmo forward,  $N^2T$ .

Los pasos para el desarrollo de este algoritmo se resume:

1) Inicialización:

$$\beta_T(i) = 1 \quad 1 \leq i \leq N$$

El cual define en forma arbitraria a todas las variables  $\beta_T(i)$  como la unidad para todos los valores de  $i$ . Estos valores hacen referencia a la última observación  $O_T$  en el instante  $t=T$ .

2) Inducción:

$$\beta_t(i) = \sum_{j=1}^N a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)$$

$$t = T-1, T-2, \dots, 1 \quad 1 \leq i \leq N$$

Donde se muestra cual es la probabilidad de poder estar en el estado  $S_i$  en el instante  $t$ , y tener en cuenta la secuencia de observaciones posteriores desde el tiempo  $t+1$  hasta  $T$ . Para esto, se tiene que considerar todos los posibles estados  $S_j$  en instantes posteriores  $t+1$ , teniendo en cuenta las transiciones desde el estado  $S_i$  al estado  $S_j$  (el termino  $a_{ij}$ ), como también las observaciones en el estado  $j$  (el termino  $b_j(O_{t+1})$ ). Finalmente, la secuencia restante parcial de observaciones que se tenían hasta el estado  $j$  los cuales están contenidos en el término de la variable Backward  $\beta_{t+1}(j)$ .

El funcionamiento de este algoritmo se observa en la Figura 14 Trellis.

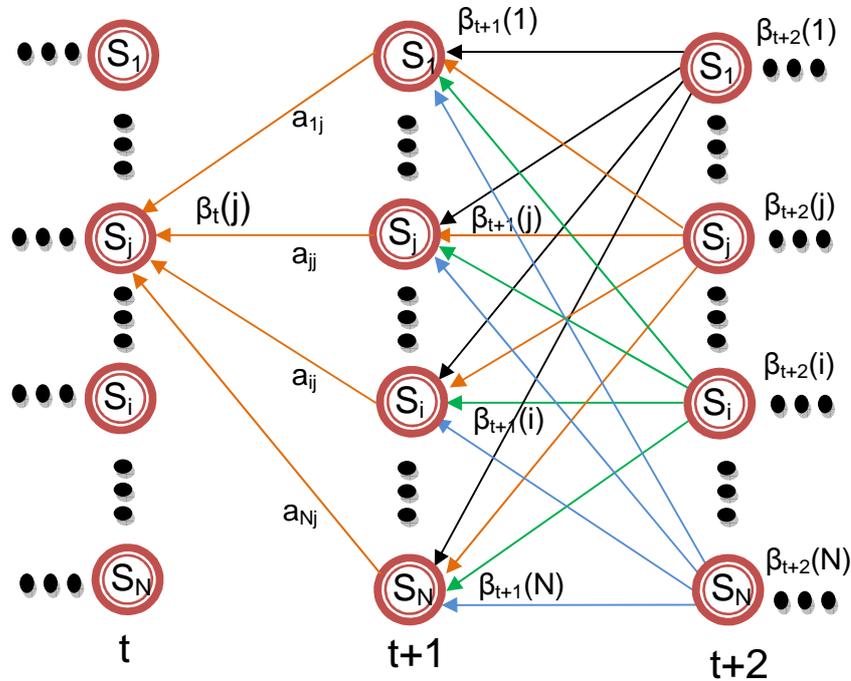


Figura 14. Proceso de cálculo de la variable Backward.

Un ejemplo de la solución de este problema es utilizado para aplicaciones que contienen varios modelos  $\lambda$ 's de HMM, y en la cual se deba seleccionar aquel modelo que mejor se adapte (que sea más probable) a una secuencia de observaciones  $O$  para efectos de reconocimiento. También calcula la probabilidad de de que ocurra un evento en un instante  $t+1$  en un estado específico, teniendo en cuenta solamente las observaciones pasadas hasta un instante  $t$ .

**3.4.2. Problema 2: Decodificación.** Dado una secuencia de observaciones  $O = O_1, O_2 \dots O_T$  (adquiridas del mundo real o simuladas), y dado un modelo definido de HMM con sus parámetros de medida probabilística  $\lambda = (A, B, \pi)$ , ¿cómo se puede detectar cual es la secuencia  $Q = (Q_1 Q_2 \dots Q_T)$  de estados ocultos más óptima que explique la secuencia de observaciones dadas  $O$ ?

Este problema se le denomina **Problema de decodificación** e intenta descubrir la parte oculta del modelo, es decir, encontrar la secuencia más probable de estados que

hayan producidos el conjunto de observaciones  $\mathbf{O}$ . Es claro, que al manejar sistemas de carácter aleatorio, es complejo encontrar una secuencia correcta y más bien se intenta encontrar una secuencia optima que mejor represente las observaciones.

A diferencia del problema 1 mencionado anteriormente, el problema de decodificación no tiene una solución exacta que defina una única secuencia de estados. Para esto existen varios posibles criterios de optimización que encuentre la secuencia más probable de acuerdo a la secuencia de observaciones dadas. La dificultad recae en definir cuál es la secuencia de estados más óptima. Una técnica formal conocida para encontrar la mejor trayectoria de secuencia de estados, se conoce como **algoritmo de Viterbi**, la cual se basa en métodos de programación dinámica.

Para encontrar la secuencia más probable de estados ocultos  $Q=(Q_1 Q_2 \dots Q_T)$ , para la secuencia de observaciones dadas  $O= O_1, O_2 \dots O_T$ , se necesita definir las siguiente nueva variable:

$$\delta_t(i) = \max_{q_1 q_2 \dots q_{t-1}} P(q_1 q_2 \dots q_t = Si, O_1 O_2 \dots O_t | \lambda)$$

Donde  $\delta_t(i)$ , es la máxima probabilidad que representa una secuencia de estados ocultos (mejor trayectoria) hasta el instante  $t$ , terminando en el estado  $q_t = S_i$ , teniendo en cuenta las primeras  $t$  observaciones. Como puede observarse, para cada instante  $t$  existe  $N$  valores  $\delta_t(i)$ , lo cual representa  $N$  secuencias de estados con máximas probabilidades (mejores trayectorias), que entran a participar a ser la secuencia más probable hasta que termine toda la secuencia de observaciones  $\mathbf{O}$ . La variable  $\delta_t(i)$  solo guarda valores de probabilidad para cada estado  $i$  y para cada instante  $t$ , lo que hace necesario además, definir una nueva variable  $\psi_t(j)$  que almacene la secuencia de estado de todas las  $N$  trayectorias. Posteriormente esta variable será utilizada para recuperar la secuencia de estados solicitada después de haberse definido cual de las  $N$  trayectoria es la más probable.

El desarrollo de este algoritmo para la construcción de la mejor secuencia de estados se resume en 4 pasos:

1) Iniciación

$$\delta_1(i) = \pi_i b_i(O_1) \quad 1 \leq i \leq N$$

$$\psi_1(i) = 0$$

El hecho que  $\psi_1(i)$  sea igual a 0 significa que no existe un estado que haya transitado al estado  $i$  por el hecho de que se está analizando para la primera observación y no existe tal estado anterior.

2) Recursión:

$$\delta_t(j) = \max_{1 \leq i \leq N} [\delta_{t-1}(i) a_{ij}] b_j(O_t)$$

$$\psi_t(j) = \operatorname{argmax}_{1 \leq i \leq N} [\delta_{t-1}(i) a_{ij}]$$

Donde  $2 \leq t \leq T$  y  $1 \leq j \leq N$

El valor  $\delta_t(j)$  para tiempos superiores a  $t=1$ , es obtenido a través de la mejor conexión (la más probable) que existe entre alguno de los estados  $i$  del instante anterior y el estado  $j$  en el instante actual (definido por el termino  $\max_{1 \leq i \leq N} [\delta_{t-1}(i) a_{ij}]$ ). Posteriormente este valor es multiplicado por la probabilidad de obtener la observación  $O_t$  en el estado  $j$ . En la Figura 15 se observa el proceso de recursión del algoritmo de Viterbi.

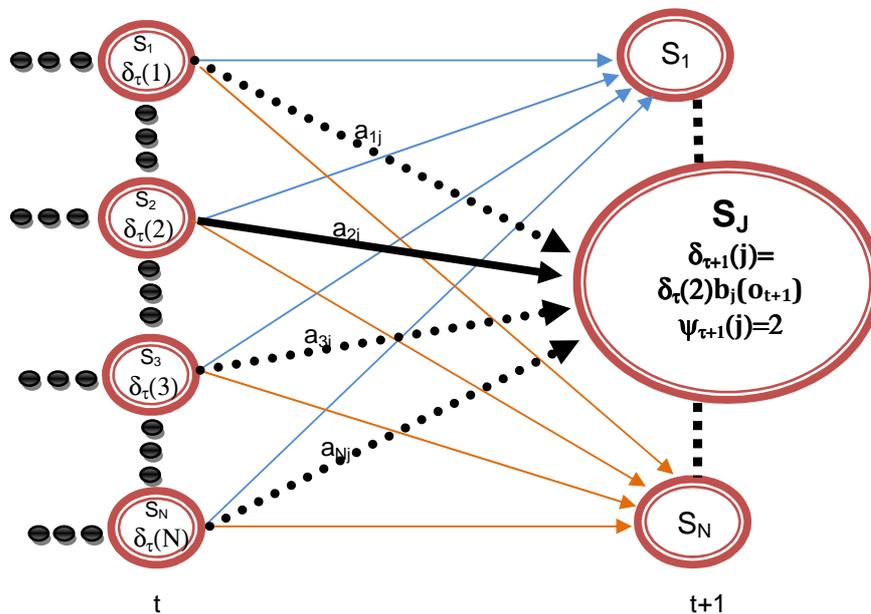


Figura 15. Ilustración del proceso de Viterbi para hallar la secuencia más probable.

El hecho de haber seleccionado la conexión mas solida del estado  $i$  al estado  $j$ , nos da una idea de cuál fue el estado anterior  $i$  en el instante  $t-1$ . Los argumentos de estos estados anteriores (representado por  $\operatorname{argmax}_{1 \leq i \leq N} [\delta_{t-1}(i) a_{ij}]$ ) son almacenados en el arreglo  $\psi_t(j)$  para cada instante  $t$  (representando el estado en el instante anterior  $t-1$ ) y cada estado  $j$ .

### 3) Finalización

$$P^* = \max_{1 \leq i \leq N} [\delta_T(i)]$$

Donde  $P^*$  es el valor de máxima probabilidad entre todas las  $N$  términos  $\delta_T(i)$  en el instante  $T$  (última observación), y en la que intrínsecamente selecciona la trayectoria más probable entre la  $N$  existentes. A partir de aquí se podría reconstruir la secuencia de estados.

Con la expresión:

$$q_T^* = \operatorname{argmax}_{1 \leq i \leq N} [\delta_T(i)]$$

Se selecciona el ultimo estado  $i$  (argumento) en el instante  $T$  el cual se obtuvo de la probabilidad  $P^*$  de todos los  $N$  términos  $\delta_T(i)$ . Teniéndose este ultimo estado se procede a la búsqueda de la secuencia de estados anteriores a este que fue representado por  $P^*$ .

### 4) Reconstrucción de la secuencia de estados (Path Backtraking):

$$q_t^* = \psi_{t+1}(q_{t+1}^*) \quad t = T - 1, T - 2, \dots, 1 \quad y \quad 1 \leq q_{t+1}^* \leq N$$

Se hace uso de la matriz  $\psi_t(j)$  para la reconstrucción de la secuencia de estados, recordándose que este almacena en el instante  $t$  argumentos de los estados de un instante anterior  $t-1$ . En la Figura 16 se observa este proceso en donde se selecciona la mejor ruta de las  $N$  mejores del proceso de recursión.

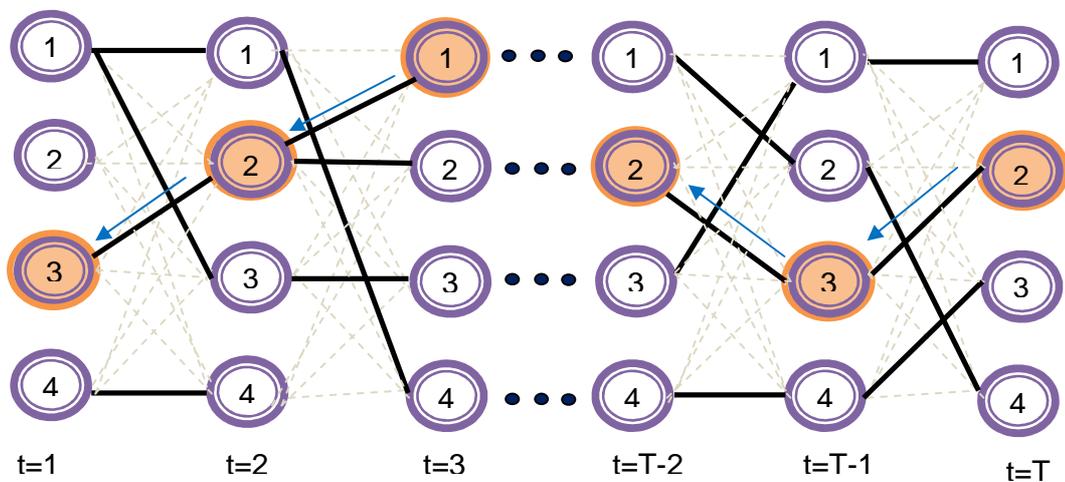


Figura 16. Proceso de Reconstrucción (Path Backtracking) de secuencia de estados.

**3.4.3. Problema 3: Entrenamiento.** Dado una secuencia de observaciones  $O = O_1, O_2, \dots, O_T$  (adquiridas del mundo real o simuladas), ¿cómo se puede ajustar los parámetros de medida probabilística  $\lambda = (A, B, \pi)$ , para maximizar  $P(O|\lambda)$ , entendiéndose como la probabilidad de que este modelo represente la secuencia de observaciones dada?

A este problema se le conoce como *Problema de Entrenamiento*, pues a partir de una secuencia  $O$  de observaciones, se busca optimizar los parámetros probabilísticos que mejor describan las observaciones dadas. La secuencia de observaciones es llamada *secuencia de entrenamiento*, y por lo general, estas secuencias de observaciones son tomadas de bases de datos representativas de fenómenos del mundo real.

No se conoce ninguna forma para resolver analíticamente para un máximo global de los parámetros del modelo. Sin embargo, hay todavía una forma para maximizar localmente la probabilidad  $P(O|\lambda)$ , usando un procedimiento iterativo llamado método de Baum-Welch, también conocido como el método EM (Expectation- Maximization).

Para describir el procedimiento de re-estimación (actualización iterativa y mejora) de los parámetros del HMM se define la siguiente variable:

$$\xi_t(i, j) = P(q_t = S_i, q_{t+1} = S_j | O, \lambda)$$

Donde representa la probabilidad de estar en el estado oculto  $S_i$  en un instante  $t$ , y estar en el estado  $S_j$  en un instante  $t + 1$ , dado el modelo y la secuencia de observaciones. Lo interesante de este concepto, es que las probabilidades de un estado  $S_i$  a un estado  $S_j$  combina las probabilidades de haber llegado al estado  $S_i$  con la variable forward y de llegar al estado  $S_j$  con la variable backward, teniendo en cuenta únicamente la conexión entre estos dos estados en un instante  $t$  dado. Esto quiere decir que dicha probabilidad se calcula sobre la posición ideal en la cual una transición debería darse en un instante determinado de acuerdo a las observaciones obtenidas  $O$ .

Los valores de esta variable  $\xi_t(i, j)$ , puede ser computado como:

$$\xi_t(i, j) = \frac{\alpha_t(i)\alpha_{ij}b_j(O_{t+1})\beta_{t+1}(j)}{P(O|\lambda)}$$

$$\xi_t(i, j) = \frac{\alpha_t(i)\alpha_{ij}b_j(O_{t+1})\beta_{t+1}(j)}{\sum_{i=1}^N \sum_{j=1}^N \alpha_t(i)\alpha_{ij}b_j(O_{t+1})\beta_{t+1}(j)}$$

En el cual para un instante  $t$  se define la probabilidad de cada par  $(i, j)$ , formando arreglos de dimensión  $N \times N$  para cada  $t$ . En la Figura 17 se observa el diagrama de trellis que muestra la conexión de estado expresado por la variable  $\xi_t(i, j)$ .

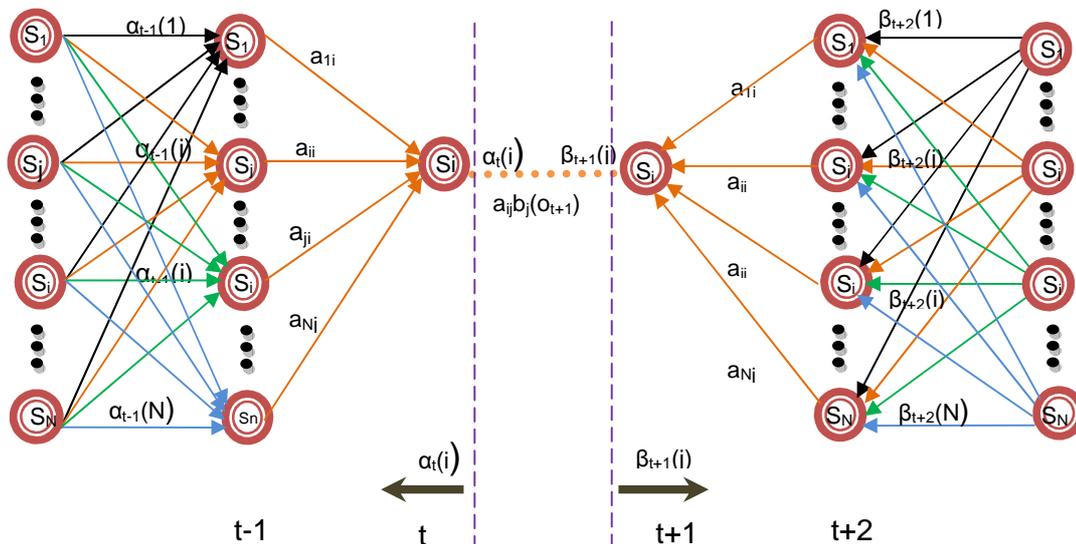


Figura 17. Proceso de entrenamiento con conexión de estados.

Se define además una nueva variable  $\gamma_t(i)$  que se define como sigue:

$$\gamma_t(i) = P(q_t = S_i | O, \lambda)$$

Donde  $\gamma_t(i)$  es la probabilidad de estar en el estado oculto  $S_i$ , en el instante  $t$ , dadas la secuencia de observaciones y los parámetros del modelo. Esta vez la probabilidad no se concentra sobre la conexión de dos estado  $(i,j)$  en un instante dado sino que tiene en cuenta la ocurrencia de un estado en un instante específico dentro de toda la secuencia de estados y acordes a las observaciones  $O$ .

Se pueden relacionar las dos variables anteriores, teniendo en cuenta que la variable  $\xi_t(i,j)$  representa la probabilidad de pasar de un estado  $i$  a uno  $j$ , mientras que la variable  $\gamma_t(i)$  solo representa la probabilidad de estar en el estado  $i$ . Si sumásemos todo los termino  $j$  del par  $(i, j)$  en un instante  $t$  se tendría como resultado, la probabilidad de estar en el estado  $i$  acorde a las observaciones y el modelo.

$$\gamma_t(i) = \sum_{j=1}^N \xi_t(i,j)$$

Estas variables son indispensables para calcular que tan frecuente se produce una transición desde un estado específico a cualquier estado, y que tan frecuente se produce una transición entre dos estados específicos.

Para el cálculo del número esperado de veces que un estado  $S_i$  es visitado se hace la siguiente suma en el tiempo sobre la variable  $\gamma_t(i)$ :

$$\text{Numero esperado de transiciones desde } S_i = \sum_{t=1}^{T-1} \gamma_t(i)$$

Igualmente el número esperado de observaciones desde  $S_i$  a  $S_j$  se calcula haciendo la suma en el tiempo de la variable  $\xi_t(i,j)$ :

$$\text{Numero esperado de transiciones desde } S_i \text{ a } S_j = \sum_{t=1}^{T-1} \xi_t(i,j)$$

Usando las expresiones anteriores se pueden calcular los parámetros de la estructura  $\lambda$  del HMM así:

$$\bar{\pi}_i = \text{numero de veces en que el estado } S_i \text{ comienza en } (t = 1) = \gamma_1(i)$$

$$a_{ij} = \frac{\text{Número esperado de transiciones del estado } S_i \text{ al } S_j}{\text{Número esperado de transiciones desde el estado } S_i} = \frac{\sum_{t=1}^{T-1} \xi_t(i, j)}{\sum_{t=1}^{T-1} \gamma_t(i)}$$

$$\bar{b}_{j(k)} = \frac{\text{Número esperado de veces en estado } S_i \text{ y observar el símbolo } v_k}{\text{Número esperado de veces en el estado } S_i} = \frac{\sum_{t=1}^{T-1} \gamma_t(i)}{\sum_{t=1}^T \gamma_t(i)}$$

La reestimación de estas formulas se guían de una función que maximiza la probabilidad de que el modelo del HMM presente el proceso dado por las observaciones  $\mathbf{O}$ . La función:

$$Q(\lambda, \bar{\lambda}) = \sum_Q P(Q|O, \lambda) \log [P(O, Q|\bar{\lambda})]$$

Representa una medida de distancia que dice que tanto el  $\bar{\lambda}$  reestimado en la nueva iteración mejora al modelo  $\lambda$  de la iteración anterior. Siempre se procura que la mejoría sea máxima tal que:

$$\max_{\bar{\lambda}} [Q(\lambda, \bar{\lambda})] \Rightarrow P(O|\bar{\lambda}) \geq P(O|\lambda)$$

Obtener el  $\bar{\lambda}$  optimo establece el conocimiento adquirido en su etapa de aprendizaje con respecto a las observaciones  $\mathbf{O}$ .

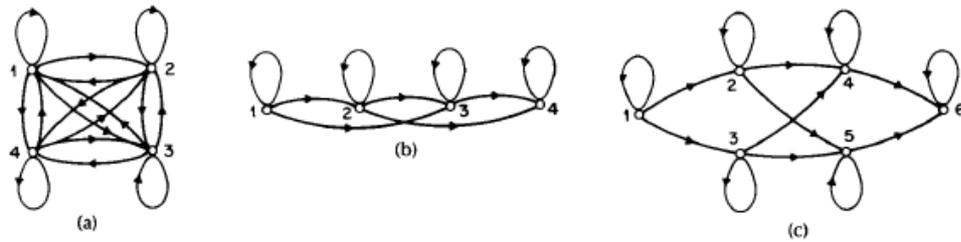
### 3.5. TOPOLOGÍAS DE LAS HMM's<sup>19</sup>

La forma en que se conectan probabilísticamente los estados de un HMM es dependiente de la naturaleza de la aplicación y de las señales que emite. En la Figura 18 se muestran las topologías más utilizadas. Entre ellas están las ergódicas (Figura 18.a) donde se permite la conexión sin restricción entre todos los estados; las de izquierda-derecha (left-right) (Figura 18.b) donde se permite transiciones de avance hacia la derecha en pasos de estados sin opción a retorno; y las de izquierda –

---

<sup>19</sup> Ibid., p.266.

derecha de conexión paralela (Figura 18.c) que permite el avance de transiciones hacia la derecha y alternarse en las diferentes trayectorias paralelas sin opción a retorno.



**Figura 18. Distintas topologías de HMM.**

Las propiedades de topología de izquierda a derecha se acomodan más a las aplicaciones en donde se presentan señales continuas cuyas propiedades cambian a través del tiempo. Estas aplicaciones tienen la suposición subyacente que la secuencia de eventos se pueden relacionar directamente a la secuencia de estados que avanzan o permanecen en el mismo estado a medida que incrementa el tiempo. Cualquier tipo de señal continua obtenida en transductores en diferentes variables de un proceso puede ser representado por esta topología.

Por otro lado las ergódicas obedecen más a cambios de estados en donde es posible en tiempos posteriores repetir ciertos eventos, caso que las de izquierda a derecha no pueden. Por lo general se implementa a aplicaciones donde las observaciones son de tipo discreto.

Las anteriores topologías aunque sean más utilizadas no restringe la existencia de otras topologías con conexiones de estados que depende exclusivamente de una aplicación particular. Por ejemplo si se aplica una topología ergódica a un problema específico, puede que dentro de esta no se necesitan ciertas transiciones entre estados haciendo que estas conexiones sean nulas.

### 3.6. DENSIDAD DE OBSERVACIONES CONTINUAS EN HMM<sup>20</sup>

La mayoría de las aplicaciones o problemas a solucionar presentan señales de tipo continuo en los cuales se presentan cambios significativos en periodos de tiempo cortos. Por lo tanto aplicar este tipo de datos a una HMM obliga a tener en cuenta ciertos replanteamientos de su estructura.

Hasta este punto del documento, se ha tenido en cuenta que el conjunto de observaciones  $\mathbf{O}$  es discreto y por lo tanto la definición de la densidad de probabilidad  $B = \{b_j(k)\}$  es de naturaleza discreta. El dominio de elementos  $V_M$  de las observaciones están, bajo este esquema, definido prácticamente sobre una matriz de  $(N \times M)$  donde  $N$  es el número de estado y  $M$  el número del dominio de posibles observaciones discretas del sistema. Sin embargo, cuando la naturaleza de las observaciones es continua, la densidad de probabilidad de las observaciones  $\mathbf{B}$  en cada estado deja de ser una matriz de elementos discretos para convertirse en una función de densidad probabilística de tipo continuo definido para todo valor real  $R$ .

La forma más general de representar esta distribución continua es a través de una función Gaussiana de múltiples mixturas definida así:

$$b_j(\mathbf{O}) = \sum_{m=1}^M c_{jm} \mathcal{N}(\mathbf{O}, \boldsymbol{\mu}_{jm}, \mathbf{U}_{jm}) \quad 1 \leq j \leq N$$

donde  $\mathbf{O}$  es el vector de observaciones a distribuir,  $C_{jm}$ , es el coeficiente de la mixtura  $m$ -ésima en el estado  $j$ , y  $\mathcal{N}$  es la función gaussiana compuestas de una media  $\boldsymbol{\mu}_{jm}$ , y una matriz de covarianza (para observaciones multidimensionales)  $\mathbf{U}_{jm}$  para la mixtura  $m$ -ésima en el estado  $j$ .

En la Figura 19 se observa una secuencia de observaciones  $\mathbf{O}(t)$  en el tiempo presentados para tres estados A, B y C con diferentes distribuciones de observaciones

---

<sup>20</sup> Ibid., p.267.

de una sola mixtura en cada uno de ellos. Las funciones gaussianas demarcan un rango que cubre aquellas observaciones que representan una clase que se diferencia de las demás. Por lo tanto un tipo de observaciones (por ejemplo, los puntos azules) son evaluados con valores insignificantes de probabilidad por la distribución de un estado diferente (por ejemplo, la función gaussiana del estado B “verde”) por el hecho de no pertenecer a esa clase. Entre menor sea las desviaciones estándares de las funciones menor dispersión y mayor selectividad tendrán las observaciones dadas a su respectiva clase. Se nota en la Figura 19.b) que la dispersión de las funciones es más pronunciada que las funciones de la Figura 19.a) teniendo como consecuencia que algunas observaciones que pertenezcan a una clase (estado) tengan las mismas probabilidades en otras clases.

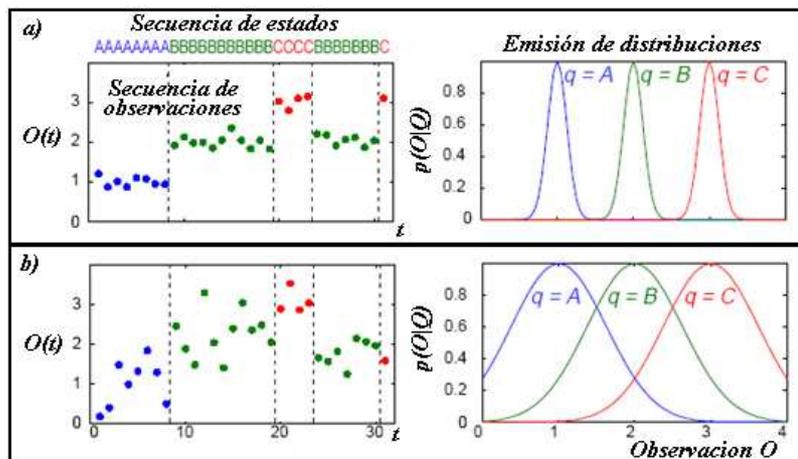


Figura 19. Observaciones Continuas y distribución de probabilidad Gaussiana.

Una combinación de mixturas se observan en la Figura 20 en donde para un estado  $S_j$  se observan 3 mixturas multiplicadas por su coeficiente  $C_{jm}$ .

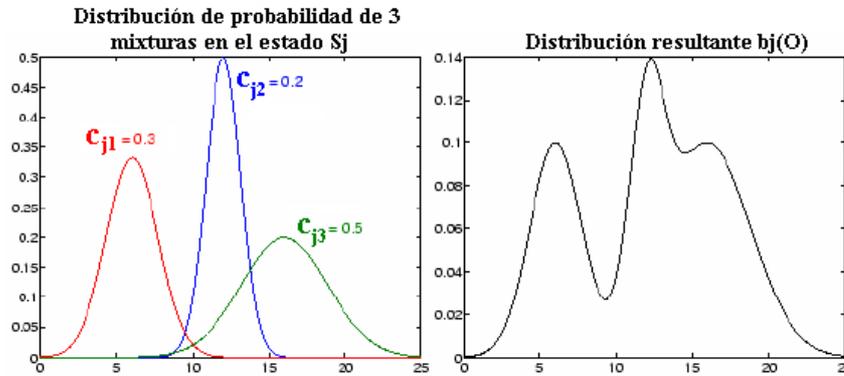


Figura 20. Mixturas Gaussianas para la densidad de probabilidad de un estado.

Los coeficientes por estado deben cumplir las siguientes propiedades:

$$\sum_{m=1}^M C_{jm} = 1 \quad \text{donde } 1 \leq j \leq N \quad \text{y } C_{jm} \geq 0$$

Cuando las observaciones  $\mathbf{O}(t)$  tengan información de más de 1 variable (por ejemplo un sistema  $n$  sensores para diversas variables como temperatura, presión, caudal, etc.) se utiliza una función gaussiana multidimensional definida como:

$$f_j(\mathbf{O}_t) = \frac{1}{(\sqrt{2\pi})^d |\Sigma_j|^{\frac{1}{2}}} * \exp \left[ -\frac{1}{2} (\mathbf{O}_t - \mu_j)^T \Sigma_j^{-1} (\mathbf{O}_t - \mu_j) \right]$$

en donde se evalúa la probabilidad de que la observación  $\mathbf{O}(t)$  de  $d$  dimensiones sea compatibles con los del estado  $j$ , por medio de una función con matriz de covarianza  $\Sigma_j$  y media  $\mu_j$  con las mismas dimensiones para el estado  $S_j$ ,

### 3.7. ALGUNAS DIFICULTADES DE IMPLEMENTACIÓN

En el tutorial de Rabiner [17] se hace una descripción muy completa sobre algunos temas que se deben tener en cuenta a la hora de hacer una implementación. Sin embargo se menciona brevemente su alcance.

**3.7.1. Escalamiento.** Este procedimiento provee una solución al problema de computación de algunos algoritmos se que se presentaron anteriormente. Para el cálculo de la variable Forward y Backward se nota la gran cantidad de multiplicaciones entre valores que son menores a la unidad. Estos hacen que el resultado final tenga una resolución tan pequeña que los recursos informáticos no provean los resultados adecuados.

**3.7.2. Secuencias de múltiples observaciones.** Provee soluciones formales a la etapa de reestimación de parámetros del modelos cuando se presentan múltiples secuencia de observaciones, siempre buscando que la probabilidad de que el modelo  $\lambda$  representa tales tipos de datos sea siempre máxima.

**3.7.3. Estimación inicial de los parámetros del HMM.** Cuando se somete el HMM a la etapa de entrenamiento es importante asignar desde un principio valores estimados de los parámetros para facilitar el proceso de convergencia del proceso. Por lo general cualquier valor de la matriz de transición  $A$  y la matriz inicial  $\pi$  es adecuado mientras que la selección de los parámetros de la densidad  $b_j(O)$  se deben seleccionar cuidadosamente por métodos de segmentación manual de los datos o utilizando el algoritmo de K- medias para encontrar las clases dentro de los datos.

**3.7.4. Efectos de no tener datos de entrenamiento suficientes.** Cuando no se dispone de datos suficientes de entrenamiento, se corre el riesgo de no obtener las estimaciones óptimas en los parámetros. Como soluciones parciales, se tiene el de

incrementar la longitud de los datos, reducir el tamaño del modelos (número de estados, número de símbolos por estado, etc.). Sin embargo en el documento [17] se provee de soluciones formales a este problema.

**3.7.5. Selección del modelo.** Dirige la atención a la selección de las estructura en cuanto al tipo de modelo (topología de estados), tamaño de la estructura (número de estados) y la selección de las observaciones (función de densidad probabilística continua o discreta, de una o varias mixturas, etc.). La selección de estos parámetros no estaba basada bajo lineamientos teóricos, por lo cual es muy dependiente del tipo de señal emitida por el sistema y a la naturaleza de la aplicación.

### RESUMEN CADENAS OCULTAS DE MARKOV

Maquina de estados aleatorios fundamentado en las propiedades de una Cadena de Markov (no oculta)

- **ELEMENTOS:** 1) Numero de estados (N); 2) Numero de observación en los estados (M); 3) Distribución de probabilidad entre estados (A); 4) Distribución de probabilidad de observaciones  $B = \{b_j(k)\}$ ; 5) Matriz de distribución de estado inicial  $\pi$ .

- **PROBLEMAS A RESOLVER POR HMM:**

**1) Evaluación:** Resuelve el cálculo de la probabilidad que un modelo definido de HMM  $\lambda$  genere un conjunto de observaciones  $O(t)$ . **Solución:** Variable Forward  $\alpha(i)$  y variable Backward  $\beta(i)$  para calculo de  $P(O|\lambda)$ .

**2) Decodificación:** Consiste en hallar la secuencia de estados ocultos que mejor represente un conjunto de observaciones  $O(t)$  dado un modelo definido de HMM  $\lambda$ . **Solución:** Cálculo de mejor trayectoria hasta un instante  $t$  con la variable  $\delta(t)$  del algoritmo de Viterbi junto con el procedimiento de backtracking.

**3) Entrenamiento:** Consiste en hallar los parámetros óptimos del modelo de HMM  $\lambda$  a partir de un conjunto de observaciones dadas  $O$ . **Solución:** Algoritmo de Baum-Welch como método iterativo en el ajuste de parámetros a partir de las concurrencias de los eventos en los estados.

- **TOPOLOGÍAS DE HMM's:** Las más utilizadas son la ergódicas y la de izquierda derecha, aunque se permiten múltiples combinaciones que se adaptan a la naturaleza de la aplicación.
- **DENSIDAD DE OBSERVACIONES CONTINUAS:** la mayoría de aplicaciones emiten señales de tipo continuo por lo que la matriz B de observaciones discretas pasan de ser una matriz a una función de densidad (gausiana) de múltiples mixturas y para múltiples dimensiones. La nueva distribución de las observaciones en los estados es:

$$b_j(O) = \sum_{m=1}^M c_{jm} \mathcal{N}(O, \mu_{jm}, U_{jm}) \quad 1 \leq j \leq N$$

- **DIFICULTADES DE IMPLEMENTACIÓN:** algunos detalles a tener en cuenta a la hora de implementar HMM.

**1) Escalamiento:** Soluciona el problema de resolución computacional ocasionado por la excesiva multiplicación de números menor que 1 haciendo los resultados insignificantes y erróneos.

**2) Secuencias de múltiples observaciones:** Somete a reestimación el modelo de HMM cuando existen múltiples secuencia de observaciones de una misma clase.

**3) Estimación inicial de parámetros:** Establece que la selección de los parámetros A,B y  $\pi$  antes del entrenamiento contribuye a la rápida convergencia del máximo (local o global). B es el parámetro más relevante que sensibiliza el éxito de la convergencia.

**4) Insuficiencia en datos de entrenamiento:** Cuando no se dispone de datos suficientes de entrenamiento, se corre el riesgo de no obtener las estimaciones óptimas en los parámetros.

**5) Selección del modelo  $\lambda$ :** Define el número de estados, topología y selección de las observaciones que serán sometidas al HMM. Se establece que su selección depende de la naturaleza de la aplicación.

### 4. APLICACIONES DE LAS CADENAS OCULTAS DE MARKOV

Las aplicaciones que se presentan en este capítulo fueron seleccionados teniendo en cuenta que los contextos de implementación fueran diferentes entre sí con el fin de mostrar la flexibilidad y los diversos ambientes en los que HMM pueden ser aplicados.

Se seleccionaron 4 aplicaciones entre ellos:

**1) Reconocimiento de Voz:** Por ser esta aplicación la primera en implementarse en los años 80's, se analiza en forma general el proceso de reconocimiento de voz, el modelamiento de las señales acústicas de una o varias palabras.

**2) Reconocimiento de lugares con Robots móviles:** Se analiza el proceso de reconocer lugares como un proceso temporal en ambientes interiores y exteriores por parte de robots móviles diseñados para tales condiciones.

**3) Técnicas predictivas de actividades de mantenimiento en la industrial: diagnóstico y pronóstico:** Muestra la necesidad de diagnosticar fallas de un sistema de poca accesibilidad y de pronosticar el tiempo útil de vida restante del equipo en análisis antes de que una falla inminente se presente.

**4) Aprendizaje y reconocimiento de patrones de movimiento corporal en seres humanos:** Se describe un modelo de comunicación mimética entre un ser humano y robot humanoide en el cual este último reconoce los patrones de movimiento ejecutados por una persona para ejecutar un movimiento de reacción o respuesta al estímulo dado.

Un esquema general para analizar cada aplicación es el siguiente:

- 1) Descripción y naturaleza del problema para justificar el uso de HMM's.
- 2) Identificación de los elementos del HMM en la aplicación.
- 3) Entrenamiento y funcionamiento del HMM dentro de la aplicación.

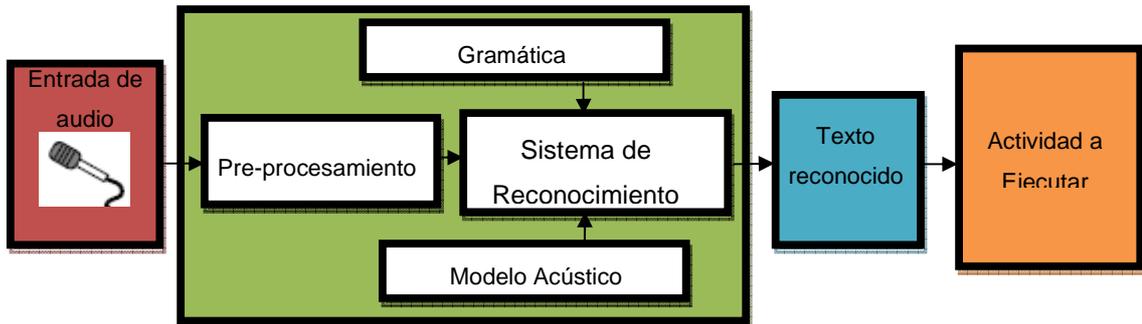
### 4.1. RECONOCIMIENTO AUTOMÁTICO DE VOZ

La forma de comunicación más cómoda, rápida y eficaz que puede existir entre seres humanos es a través del habla. Debido a la comodidad que nos produce, existe un deseo por interactuar de esta manera con las máquinas a través de comunicación hablada.

La habilidad que tendría una máquina de convertir señales de voz automáticamente en texto se torna interesante en un amplio rango de contextos, entre ellos, la interacción con computadores (dictados), transmisión de órdenes de control sobre un sistema, operaciones de manos libres (hand-free), búsqueda de información de una base de datos, traducción automática de lenguaje, ejecución de comandos de control mediante voz para personas discapacitadas, en fin, gran variedad de interfaces hombre-máquina controlados por la voz para sustituir en ciertas aplicaciones a los interfaces tradicionales basados en teclados, paneles y dispositivos similares.

En las siguientes secciones relativas a esta aplicación se tratarán aspectos relacionados a la naturaleza del reconocimiento de voz en donde se manifiestan las dificultades y factores que influyen en la producción y reconocimiento de voz; la identificación de esta aplicación como un proceso doblemente estocástico para justificar el uso de HMM's; criterios de selección de la estructura  $\lambda$  (Estados, Topología, Observaciones y funciones de distribución de probabilidad) del HMM; el proceso de entrenamiento y por último el proceso de reconocimiento de voz.

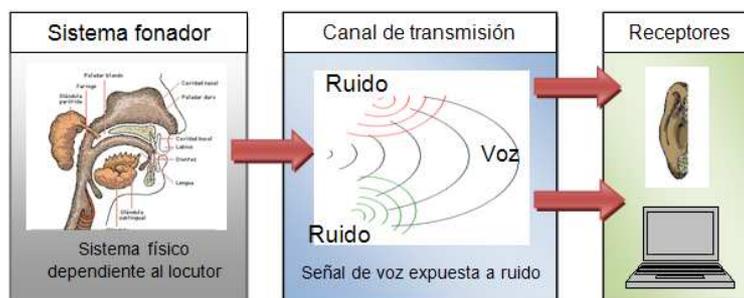
**4.1.1. Naturaleza de la Generación y el Reconocimiento de Voz.** El esquema general de un reconocedor de voz para la ejecución de una tarea particular integra múltiples etapas como se muestra en la Figura 21 que transforman una entrada de audio de modo natural y lo trasladan a texto conocido.



**Figura 21. Estructura General de un Reconocedor Automático de Voz.**

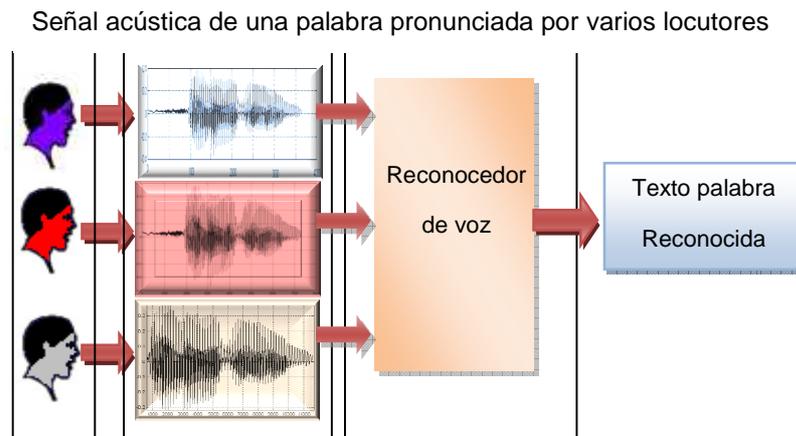
La obtención apropiada del texto a reconocer por parte del sistema de reconocimiento depende de cómo éste interprete la señal recogida en la entrada de audio. Por lo general, la interpretación es apoyada de un modelo acústico que correlaciona las ondas sonoras con unidades de texto identificables, y un modelo gramatical que las organiza para que tengan un significado coherente. La dificultad que se presenta en tales actividades ejecutadas por estos modelos radica en que dicha interpretación es afectada por varios factores de incertidumbre y variabilidad encontradas en la información que llega al receptor.

En la Figura 22 se observa la presencia de algunos de estos factores que el receptor humano debe enfrentar en el reconocimiento y que de manera similar un sistema artificial debe ser sometido.



**Figura 22. Proceso de pronunciación, transmisión y recepción de la voz.**

Uno de los factores está muy relacionado al sistema fonador humano, responsable en la pronunciación los diferentes sonidos, el cual presenta casos de variabilidad *interlocutor* e *intraocutor*. Debido a que el sistema articulatorio fonador es un sistema biológico único para cada individuo, el sonido generado será fuertemente dependiente del hablante actual. La **variabilidad intraocutor** describe que una persona no pronuncia siempre de la misma forma una palabra específica debido a causas influidas, por ejemplo, al estado de ánimo en que se encuentre el locutor, la velocidad y el acento con que se haga la pronunciación, entre otras. Por otro lado, la **variabilidad interlocutor** resalta que no importa quién pronuncie un sonido específico (pitch, genero) las ondas emitidas serán entendidas por quien la escucha y deben ser entendidas igualmente por el sistema artificial. En la Figura 23 se ilustra la situación en donde diferentes locutores generan diferentes formas de onda para una misma palabra u oración y el sistema debe entenderlas como si fuera un mismo objeto.



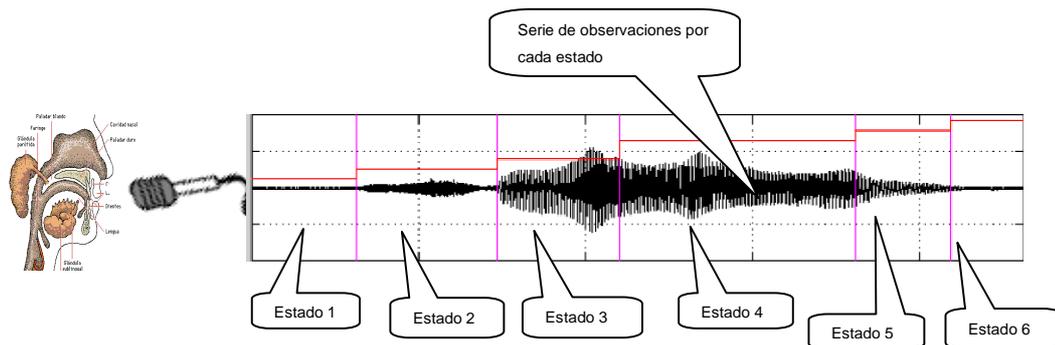
**Figura 23. Reconocimiento de un mismo texto para diferentes locutores.**

Otro factor que dificulta el reconocimiento de voz para una maquina, es el influido por el entorno que participa como canal en la transmisión de la voz. En este caso, un conjunto de palabras pronunciadas en un entorno con sonidos adicionales ajenos al mensaje transmitido, o con algunos elementos de distorsión (ecos, reverberaciones),

introducen ruido y perturbaciones dentro de la señal emitida dificultando aun más el proceso de reconocimiento.

A pesar de estas alteraciones de la señal de voz, los receptores humanos tienen la capacidad de abstraer el mensaje transmitido en forma clara y bien entendida, situación que en semejantes condiciones debe el sistema artificial de reconocimiento obtener los mismos resultados.

El modelo acústico y gramatical del sistema de reconocimiento puede ser representado por modelos HMM's, que gracias a su flexibilidad toleran cada uno de los factores mencionados. Bajo este esquema, la generación de voz puede ser vista como un proceso doblemente estocástico. Los dos procesos (observable y no observable) estocásticos en la generación de voz se ilustran en la Figura 24 para una palabra u oración cualquiera.



**Figura 24. Identificación de Doble Proceso Estocástico en Generación de Voz.**

En el proceso oculto (no observable), la voz es producida por los lentos movimientos de órgano articulatorio que siguiendo una secuencia de diferentes posiciones, produce una serie de sonidos que forman la señal de voz. Cada posición articulatoria puede ser representada por un estado de diferente y variable duración, y la transición entre diferentes posiciones puede ser representado por la matriz  $A = \{a_{ij}\}$ . Por otro lado, en el proceso observable, las observaciones obtenidas por micrófono son sonidos producidos de acuerdo a la posición articulatoria (estado sistema fonador) en el que se encuentre y que puede ser representado por la función de probabilidad  $B = \{b_j(O)\}$  debido a las variaciones en la evolución de cada sonido.

Dada la alta variabilidad temporal de la señal de voz ocasionado por las dinámicas variaciones del sistema articulador, se pensaría que un gran número de estados cubrirían todas las posibles clases de sonidos que podrían existir dentro de la señal de voz. Sin embargo, encontrar el número mínimo óptimo de estados que cubran sin dificultades la mayoría de sonidos es lo que reduce tiempo de procesamiento de esta señal. Por lo general, se correlaciona manualmente una variación particular en las observaciones provenientes del sistema articulador con una unidad de sonido que se determina como estado que sea conocida en nuestro lenguaje como por ejemplo, un fonema, una sílaba, una palabra, etc., dependiendo de la estructura diseñada del HMM.

A continuación se mencionan algunos criterios generales de selección de la estructura del HMM para el reconocimiento de voz los cuales cubren aspectos como la selección de la unidad de sonido representada por un solo HMM, el número de estados, la topología, la obtención de conjunto de observaciones  $\mathbf{O}$  y las funciones de densidad de probabilidad.

**4.1.2. Selección Estructura  $\lambda$  de HMM para el Reconocedor de Voz.** La selección de la estructura del HMM se define de acuerdo a algunas condiciones requeridas en la aplicación de reconocimiento de voz<sup>21</sup>, entre ellas se mencionan las siguientes:

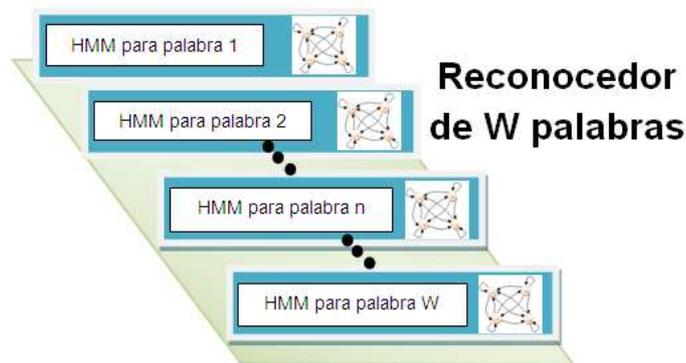
1) **Nivel de complejidad del reconocedor de voz.** En los reconocedores de **palabra aislada**, por ejemplo, deber reconocer únicamente palabras que han sido pronunciadas en forma aislada y cada una de estas palabras deben ser tratadas independientemente por el reconocedor. Los reconocedores de **palabras conectadas** requieren que las palabras en una oración sean pronunciadas en forma separada por medio de periodos artificiales de silencio. Sin embargo, contrario al reconocedor de palabras aisladas, la importancia aquí es la secuencia de palabras y no solamente cada palabra en la secuencia. Por último, están los reconocedores de **habla continua** son aun más complejos que los anteriores, debido a que deben reconocer oraciones que son pronunciadas en una manera natural sin pausas artificiales entre palabras.

---

<sup>21</sup> RIIS, Soren Kamarik, *Hidden Markov Model and Neural Networks for Speech Recognition*, 1998. Ph.D. Thesis.

2) **El tamaño del vocabulario.** La cantidad de palabras a reconocer por un sistema define la unidad de reconocimiento de voz básica denominada como SMU<sup>22</sup> (Symbolic Message Unit, por sus siglas en ingles) que en conjunto conforman el vocabulario completo. Cada una de estas unidades SMU es modelada por un solo HMM. Si el reconocimiento solo requiere hacerse sobre unas pocas palabras por lo general se utiliza cada palabra como un SMU. Si el tamaño del vocabulario es considerablemente grande es casi mandatorio elegir el SMU a una unidad más pequeña que una palabra misma, en este caso, fonemas o silabas, que concatenadas en serie forman unidades superiores a su composición<sup>23</sup>, como palabras, conjunto de palabras, frases, párrafos, etc.

En la Figura 25 se observa un conjunto de reconocimiento de palabras aisladas conformado por  $W$  elementos en donde cada una de estas SMU's corresponde a un HMM.



**Figura 25. Reconocedor de palabra aislada de  $W$  elementos.**

Definido el tamaño del vocabulario, los elementos que conforman la estructura de cada HMM se mencionan:

**4.1.2.1 Selección del número de Estados y topología.** La definición del número de estados del HMM para un reconocedor de palabra aislada cuyo SMU corresponde a una palabra según Rabiner en [17] tiene dos líneas de pensamiento. Una de ellas, es

---

<sup>22</sup> Ibid., p.3.

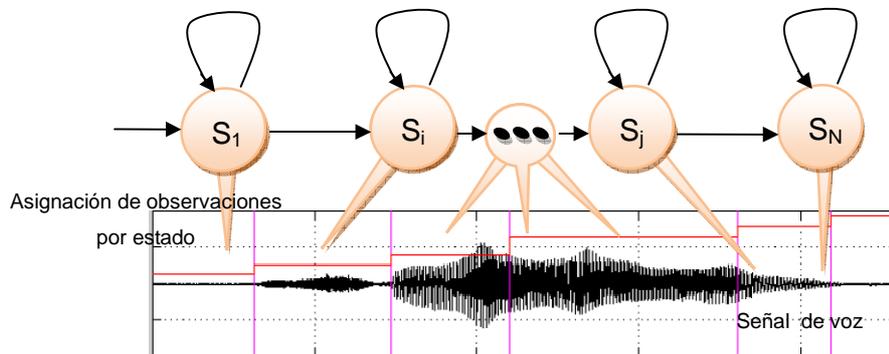
<sup>23</sup> Ibid., p.33.

la de asignar el número de estados relacionado al número de sonidos (fonemas) o silabas dentro de una palabra, que por lo regular puede contener entre 2 y 10 estados. La segunda idea, es la de asignar abruptamente a cada estado un número promedio de observaciones de una versión hablada de una palabra. La primera idea es la más utilizada, debido a que un estado en particular tiene un significado para el diseñador por el hecho de ser el fonema o silaba una entidad reconocible y un evento distinguible de otros eventos dentro de la señal.

En otros tipos de reconocedor de gran vocabulario cuyo HMM es diseñado para unidades SMU inferiores a la palabra (fonema o silaba), por ejemplo en [15], el número de estados es difícil de estipular debido a que no se tiene un conocimiento directo y empírico de las unidades identificables que podrían conformar un fonema o silaba.

Algunas veces se utilizan técnicas para la identificación previa de clases o clústeres en la señal como se hace en [17] el cual utiliza el algoritmo de K-medias (*K-means*) para la distribución de las observaciones en los diferentes estados como parámetros iniciales previos al entrenamiento. Otras técnicas con los mismos fines se mencionan al algoritmo K-NN (Nearest Neighbour), Redes neuronales de Perceptron Multicapa, Mapas auto-organizadores de Kohonen, Growing Neural Gas, entre otros que identifican clases sea bajo el esquema de aprendizaje supervisado o aprendizaje no supervisado.

Con respecto a la topología que involucra la conexión entre estados en el HMM, la más apropiada para modelar una palabra específica, es aquella conocida como topología de izquierda- a - derecha (left - to - right) como se ilustra en la Figura 26.



**Figura 26. Topología Left-to-Right de HMM utilizada en el reconocimiento de voz.**

Esta configuración explica de manera más eficiente el desarrollo del proceso natural de la pronunciación de la voz en un sentido “hacia adelante” (tal como se lee, se pronuncia), y comprende las reglas de pronunciación secuencial (léxico) de los sonidos para una palabra específica, asociando el flujo temporal de la señal de voz con los estados en una manera sencilla<sup>24</sup>. Se observa que a cada estado se le asigna un conjunto de observaciones que representa un evento discriminante dentro de la señal de voz.

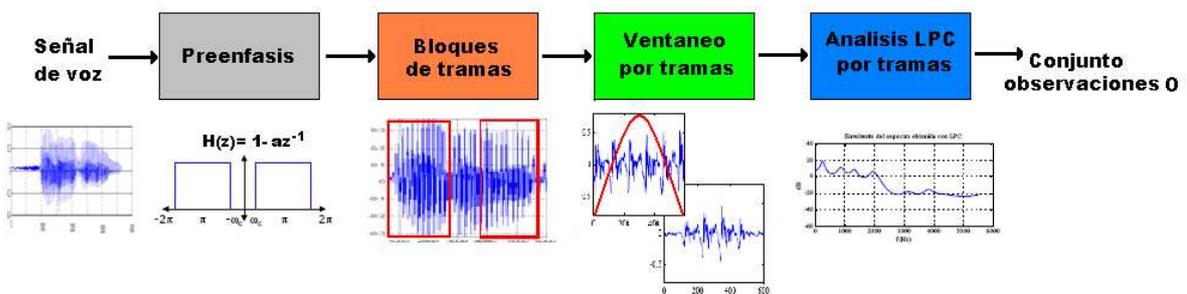
**4.1.2.2. Definición del conjunto de observaciones  $O$ .** Para utilizar un HMM en reconocimiento de voz, el vector de observaciones  $O$  a los cuales estará sometido el modelo es el resultado de una etapa de pre-procesamiento que convierte y comprime los datos crudos (*raw data*) obtenidos por el micrófono a un conjunto de datos de vectores característicos de la voz del locutor. Un conjunto de datos crudos obtenidos, por ejemplo, para una palabra pronunciada relativamente corta de aproximadamente 2 segundos, y muestreada por un computador a una frecuencia entre 8 y 16 KHz arrojaría entre 16000 y 30000 observaciones con distintas amplitudes y frecuencias. Si en forma recurrente se pensara en reconocer series de palabra u oraciones con duraciones entre 5 y 10 segundos, la cantidad de datos obtenidos sería de gran

<sup>24</sup> ABDULLA, Waleed, KASABOV, Nikola. *The concepts of Hidden Markov Model in Speech Recognition*. New Zealand, 1999. Technical Report. University of Otago. Information Science Department.

magnitud y la tarea de reconocimiento de frases serían inefficientes. Una elevada serie de observaciones para un HMM representaría problemas en eficiencia y complejidad computacional.

Como solución a esa problemática, existen métodos que comprimen los datos hasta en un 10 % extrayendo el conjunto de vectores de características (*feature vectors*) de la señal de voz, preservando información del mensaje pronunciado y permitiendo así la recuperación del conjunto original de datos. Los vectores de características son utilizados tanto en la etapa de entrenamiento como en etapa de reconocimiento del HMM. Alguno de esos métodos consisten en técnicas de análisis espectral y/o temporal como Análisis de Fourier, Codificación Predictiva Lineal (LPC: Linear Predictive Coding), Análisis de Coeficientes Cepstrales, Predicción Lineal Perceptiva, Bancos de filtros, transformadas de Wavelet, etc.

El procesamiento para la obtención del conjunto de observaciones **O** de características a partir de la señal de voz obtenida por micrófono se observa en la Figura 27.



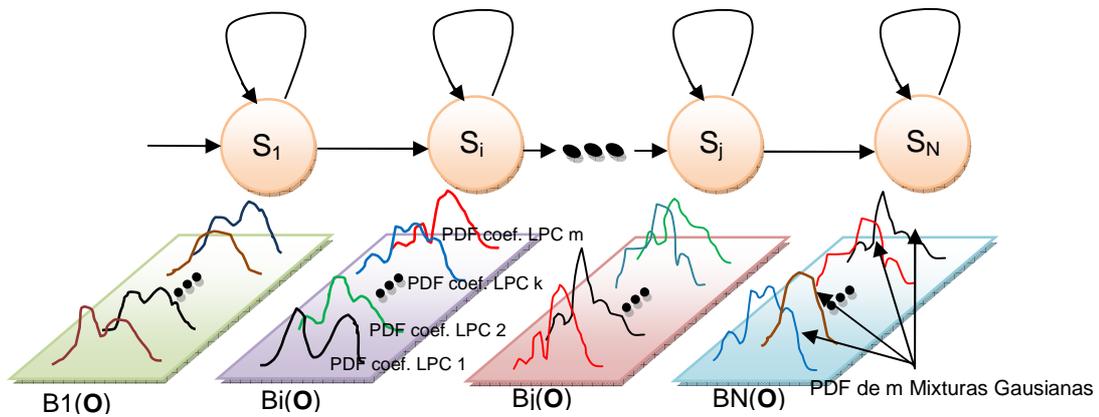
**Figura 27. Obtención conjunto de observaciones O por pre-procesamiento de señal.**

En la etapa de **Preénfasis** se aplica un filtro pasa altos de primer orden con el fin de enfatizar las frecuencias altas y allanar la señal espectralmente. En **Bloque de tramas** se secciona la señal en conjuntos de N datos (entre 200 y 300 muestras) y se avanza en pasos de M datos (menor que N, entre 50 y 100 muestras), teniendo datos traslapados entre trama y trama para mantener la dependencia temporal y abarcar toda la señal de voz a analizar. En la etapa de **Ventaneo por tramas** se multiplica cada trama obtenida en la etapa anterior por una ventana de Hamming (o Hanning) de

la misma longitud de trama  $N$  con el fin de reducir la distorsión espectral ocasionada por segmentación abrupta en tramas de la señal. Finalmente se aplica una técnica de análisis temporal o espectral, en este caso **Análisis LPC por tramas** para obtener los vectores de características de cada trama. Con este último se hace la recolección de toda la secuencia de vectores característicos de toda la señal siendo estos el conjunto de observaciones  $\mathbf{O}$  que puede ser utilizada en el entrenamiento o en el reconocimiento del HMM más probable a la palabra mencionada.

El conjunto original de datos (aquellos obtenidos por el micrófono) a pesar de su gran tamaño, es representado por una sola variable de amplitud (vector unidimensional). Sin embargo, los vectores de características obtenidas en el pre-procesamiento, incrementa el número de variables (vector multidimensional) de los datos de acuerdo al número de coeficientes LPC requeridos por el diseñador. Por lo general el orden o dimensión de los datos puede incrementarse de 10 a 30 dependiendo de la técnica de análisis. A pesar de aumentar la dimensión del vector de observaciones, esto no afecta la reducción de su cantidad.

**4.1.2.3. Relación entre las Observaciones y los Sonidos Pronunciados.** Esta relación define los elementos de la función de densidad de probabilidad (PDF por sus siglas en inglés)  $B_j(\mathbf{O})$  donde para cada estado  $j$  (fonema, sílaba o palabra) se relaciona probabilísticamente la ocurrencia de los elementos del vector de Observaciones  $\mathbf{O}$  compuestos por coeficientes LPC de la señal de voz. Por lo general, la naturaleza del conjunto de observaciones es de carácter continuo y las funciones de distribución de probabilidad  $B_j(\mathbf{O})$  por estado son definidas para variable continua, comúnmente distribuciones Gaussianas de  $n$  orden ( donde  $n$ : número de coeficientes LPC). En otros casos se utilizan densidades de símbolos de observaciones discretas, en los cuales se utilizan un cuantizador de vectores (VQ: Vector Quantization) para mapear cada vector de observación continua a un índice discreto a partir de un libro de código. En la Figura 28 se ilustra la representación de varias funciones gaussianas de distribución de probabilidad de las observaciones para cada estado.



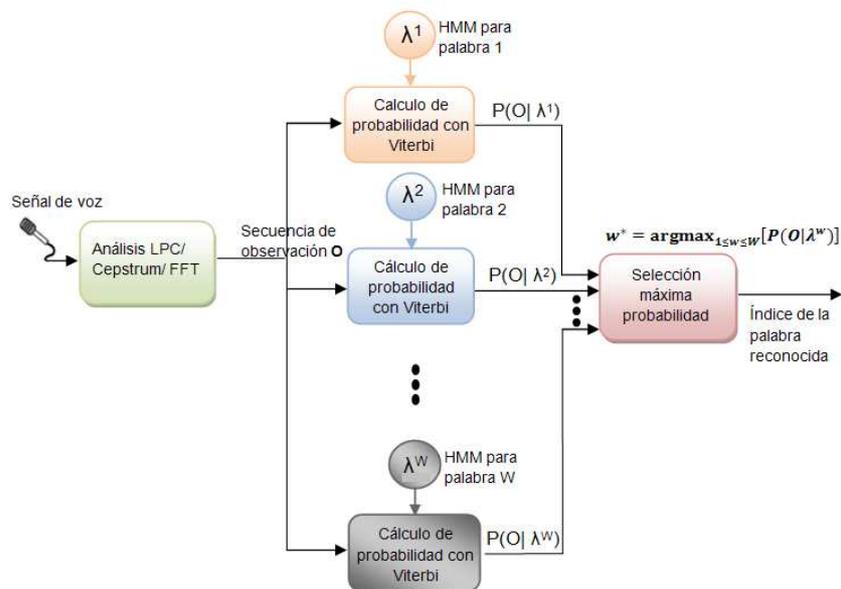
**Figura 28. Función de densidad de probabilidad de las observaciones  $O$  por Estado.**

Se observa que cada uno de los estados contiene una función de densidad de probabilidad  $B_j(O)$  compuesta por distribuciones gaussianas de múltiples mixturas para cada coeficiente LPC. La ventaja de usar la representación probabilística que ofrece HMM es que la distribución de datos dentro de cada SMU puede ser configurada por aprendizaje automático a partir de un conjunto de pronunciaci3nes de entrenamiento. Adem3s, la representaci3n probabilística es eficiente para la representaci3n de variaciones entre locutores.

**4.1.3. Entrenamiento y Reconocimiento de Voz.** El conjunto de entrenamiento para el reconocedor voz contiene datos de locutores de diferentes edades y géneros. En [17] se utiliza un set de entrenamiento consistente en 100 ocurrencias de cada palabra generado por 100 locutores (una sola pronunciaci3n de cada palabra por locutor), los cuales la mitad del conjunto son pronunciaci3nes de hombres y la otra mitad de mujeres. Se utiliz3 tambi3n otros conjuntos diferentes al de entrenamiento para evaluar las etapas de reconocimiento para cada una de las palabras entrenadas. Existen adem3s bases de datos estandarizadas que han sido creadas para permitir comparaciones bajo condiciones bien definidas. Algunos ejemplos de tales bases de datos esta el TIMIT de voces leídas de Ingles Americano; The PhoneBook database que contiene pronunciaci3nes de palabras asiladas sobre las líneas telef3nicas de

Norte América; el DARPA resource Management database que contiene oraciones leídas en un campo de manejo de recursos en el campo naval; la ARPA Wall Street Journal database que contiene pronunciaciones leídas de la revista de Wall Street de Nueva York.

En la etapa de reconocimiento de voz se busca encontrarle el sentido o significado a la señal de voz obtenida para su identificación y transformación a texto coherente. Descubrir el sentido de la señal, consiste en averiguar la serie de estados por los cuales el sistema fonador ha tenido que desarrollarse para producir una secuencia de sonidos específicos que conforman una palabra. Este problema compete a la solución del algoritmo de Viterbi el cual encuentra la secuencia óptima de estados del sistema en cada uno de los modelos HMM del vocabulario diseñado y seleccionar el más probable. Luego de haber entrenado todos los modelos HMM del vocabulario, el reconocimiento de una palabra desconocida es llevado a cabo por el proceso que se ilustra en la Figura 29.



**Figura 29. Diagrama de bloques de un reconocedor de palabras aisladas con HMM.**

La señal de voz desconocida, al igual que en la etapa de entrenamiento, se le realiza un análisis espectral con el fin de obtener la secuencia de observaciones  $\mathbf{O}$ . Acto

seguido, se calcula la probabilidad de las observaciones en todos los  $W$  modelos HMM por medio del algoritmo de Viterbi (sobre la trayectoria de estados más probables). Por último se selecciona aquel modelo con la más alta probabilidad por medio del índice que lo identifica.

### RESUMEN RECONOCIMIENTO DE VOZ CON HMM.

#### 1. Naturaleza y dificultades del reconocimiento de Voz

Permitir a un sistema artificial que reconozca una serie de ondas acústicas de voz y lo transforme en texto, teniendo en cuenta los factores de variabilidad inducidos por el locutor (intra e inter-locutor) y el canal transmisor de la voz (ruido).

#### 2. Estructura de HMM.

- **Lo que representa un HMM:** Dependiendo del tamaño del vocabulario se puede designar a un HMM (o SMU) como un fonema, sílaba, palabra, o conjunto de pequeñas palabras.
- **Topología:** Por lo general se utiliza la de Izquierda - a- Derecha (Left- to Right) por representación sencilla del flujo de voz como secuencia temporal bajo esquema "Tal como se escribe, se lee".
- **Numero de estados:** Dependiendo del SMU seleccionado se identifica los posibles sonidos conocidos dentro de esta, sean fonema, sílabas, etc. que lo conforman. Si no es posible conocerlo directamente se utiliza técnica de clusterización de datos para la identificación de clases dentro del SMU.
- **Observaciones O:** el conjunto es conformado por el resultado de una etapa de Pre-procesamiento que arroja la serie de vectores de características (feature vectors) de alguna técnica de análisis espectral o temporal, ejemplo, Coeficientes LPC, Cepstrum, FFT, etc. La distribución de probabilidad de las observaciones dentro de los estados esta compuestos por funciones gaussianas de múltiples mixturas para cada una de las variables que conforma la dimensión del vector de observaciones.

#### 3. Entrenamiento de los Modelos HMM

Se utiliza bases de datos de múltiples (sexo, edad, contexto) locutores para el ajuste de los parámetros de probabilidad de transición entre estados, distribución de observaciones por estado por medio de algoritmo E-M o Baum- Welch. Algunas bases de datos estándar son utilizadas abiertamente para permitir comparaciones después del entrenamiento, como TIMIT, PhoneBook, DARPA, ARPA, entre otras.

#### 4. Reconocimiento de Voz.

Se pre-procesa la señal de voz de prueba para obtención del vector de observaciones, se evalúa la probabilidad de todos los modelos HMM entrenados por medio de el algoritmo de Viterbi (por la ruta de estados más probable) y se elige el HMM con la más alto valor probabilístico.

### 4.2. APRENDIZAJE Y RECONOCIMIENTO DE LUGARES CON ROBOTS MÓVILES

Los robots autónomos móviles se han estado desarrollando para múltiples propósitos y su esencia principal es de la realizar tareas de navegación con capacidad para defenderse autónomamente en forma apropiada dentro de un ambiente de variadas condiciones.

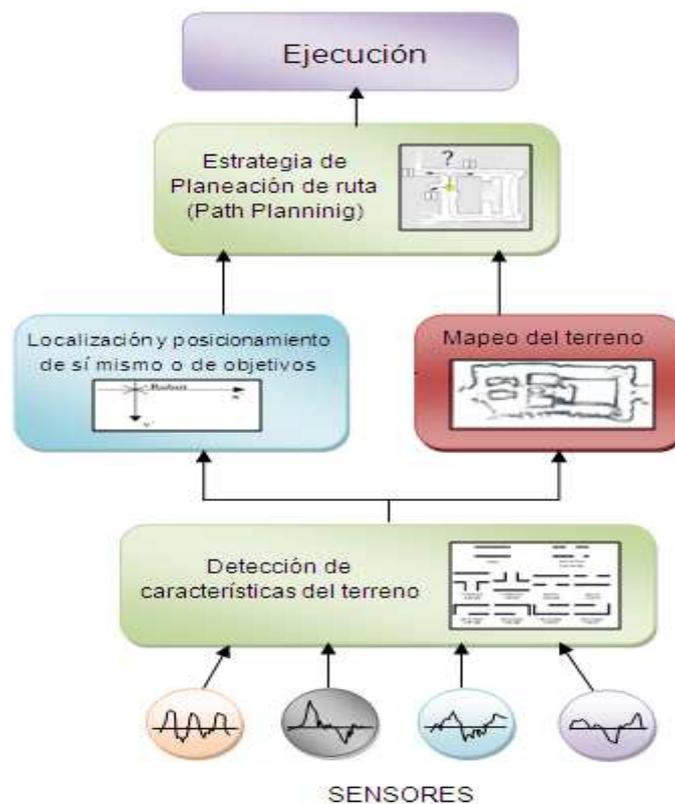
Entre las tareas principales que desempeñan estas maquinas, se mencionan por ejemplo la localización de objetos, exploración de nuevos espacios, identificación de la mejor ruta, transporte de objetos a destinos designados, mapeo para la identificación de condiciones de terrenos, y en fin una gran cantidad de aplicaciones que el lector podría imaginarse.

En las siguientes secciones se trataran aspectos relacionados a los procesos y dificultades que repercuten en el reconocimiento de lugares con robots móviles, la definición de las características de un terreno abierto o cerrado, los criterios de selección de la estructura  $\lambda$  (Estados, Topología, Observaciones y funciones de distribución de probabilidad) del HMM, el proceso de entrenamiento para algunas condiciones del terreno y por último el proceso de reconocimiento de sitios en forma autónoma.

**4.2.1. Naturaleza del reconocimiento de lugares con Robots Móviles.** El propósito de los científicos de proporcionarle la autonomía a una maquina suprime en la mayoría de los casos la necesidad de tener que asignar un operario humano para guiarla por la ruta deseada haciendo este uso de sus capacidades en la interpretación de la información relacionada al terreno. Sin embargo, en muchas ocasiones la interpretación que los operadores dan a los datos de sensores, especialmente a la combinación de varios de ellos, es compleja debido a que el ser humano no está en la capacidad de analizar tal cantidad de información en tiempo real. Por otro lado, la autonomía de un robot también provee soluciones a problemas de comunicación limitada y retardada entre el operador y el robot, como es el caso de los exploradores planetarios (*planetary rover*), donde la interacción humana está restringida, y por lo

tanto la detección de las diferentes condiciones del entorno para la construcción de la estrategia de movimiento apropiada de navegación debe ser llevada a cabo por él mismo.

El desarrollo de su autonomía integra múltiples procesos de diferentes niveles que se ejecutan en función de la complejidad de las condiciones encontradas en el ambiente que se enfrenta el robot móvil y las necesidades planteadas por los diseñadores en el cumplimiento de una tarea de navegación. En la Figura 30 se observa el esquema general de los procesos para el desarrollo de la navegación para un robot móvil.



**Figura 30. Niveles del Proceso de Navegación en Robots Móviles.**

En [27] los autores operaron manualmente robots móviles Pioneer AT y Segway RPM con sensores laser, para mapear el terreno compuesto por grama y pisos de concreto a través de las características detectadas del terreno obtenidas en sus sensores. En [26] se estudian diferentes técnicas métricas y topológicas de mapeo de sitios cerrados

(indoor environment) entre ellos, filtros de Kalman, algoritmos de LU/Minios, mapas con mallas de permanencia (occupancy grids map), entre otros., utilizando información de las características del terreno a través de los sensores. En [9] se desarrollan algoritmos en robots autónomos para la planeación y ejecución de movimiento, haciendo uso de la técnica de localización y mapeo autónomo (SLAM: Simultaneous Localisation and mapping) en ambientes con obstáculos estáticos y dinámicos utilizando información de sus sensores. En [25], [5] y [6] se tratan los problemas de planeación de ruta y ejecución de movimiento en robots móviles haciendo análisis del terreno y estableciendo las mejores rutas acordes a las representaciones obtenidas en sus sensores.

En estos y muchos otros trabajos con robots móviles, es fácil notar que la eficiencia de los diferentes niveles del proceso de navegación depende de la interpretación que brindan estas máquinas a la información obtenida por la variedad de sensores utilizados para detectar las características del terreno. Una característica del terreno puede ser definida como un lugar (pasillo, puertas, esquinas, intersecciones, etc.) o una situación (pendientes, cruce de obstáculos, desniveles, etc.) dependiendo del entorno a reconocer. La dependencia del robot a la información obtenida por los sensores requiere de métodos que lo conduzcan a una óptima interpretación de dicha información con respecto a lo que encuentre a su paso. Sin embargo, tal interpretación está sujeta a enfrentarse a algunas dificultades que tienden a obstaculizar el éxito de la misma. A continuación se presentan estos factores.

**4.2.1.1. Dificultades de interpretación de la información de sensores en Robots móviles.** De una interpretación aceptable de la información del terreno depende que los niveles superiores a este (mapeo, localización, planeación y ejecución, Figura 30) sean ejecutadas en forma adecuada, evidenciando la necesidad de encontrar una manera de relacionar la información obtenida en los sensores con las condiciones reales detectadas en el terreno. Existen gran cantidad de sensores utilizados en robots móviles, distinguiéndose aquellos de bajo costo, bajo consumo, procesamiento rápido que permite una respuesta rápida del robot a condiciones detectadas. Entre ellos se encuentran los sensores no visuales, como el ultrasonido, los infrarrojos, táctiles, etc.,

que transforma alguna cantidad medida en una señal eléctrica sencilla. Existen sensores más sofisticados, como los visuales, destacándose las cámaras de video, que a pesar de obtener información que otros sensores no pueden, el tiempo de procesamiento es lento comparado con estos últimos. Frecuentemente son utilizados los sensores no visuales por la simplicidad de sus lecturas que permiten un rápido procesamiento de la información y generar respuestas rápidas a estímulos del entorno.

En vista de lo anterior, la interpretación de la información tomada por los sensores del robot móvil está sometida a los siguientes factores:

- Ruido inherente en la lectura de sus mediciones e incertidumbre de lo que realmente se está observando por pérdida fortuita de información o ambigüedades por causa de acumulación de tal ruido.
- Sometimiento a estados de confusión inducido por problemas de aliasing perceptual [8] en donde situaciones distintas en las que se encuentra el robot y en las que deberían tomar acciones también distintas, se solapa en una misma percepción. Caso contrario se presenta cuando se produce soluciones distintas a situaciones de condiciones de terreno similares en la que deberían ser solucionadas por una acción común.
- Pobre correlación entre los datos generados por grandes cantidades de sensores utilizados en tiempo real para representar las características detectadas de terreno.

Una Cadena Oculta de Markov puede relacionar estadísticamente la información obtenida en los sensores y las características del terreno, lidiando aceptablemente los factores anteriormente mencionados. Sus parámetros representan la evolución de la información obtenida secuencialmente en el tiempo como una descripción de lo que observa a su paso a medida que el robot avanza por el terreno. En términos generales, los cambios registrados en la medición de los sensores en el tiempo manifiestan cambios que describen las condiciones del terreno.

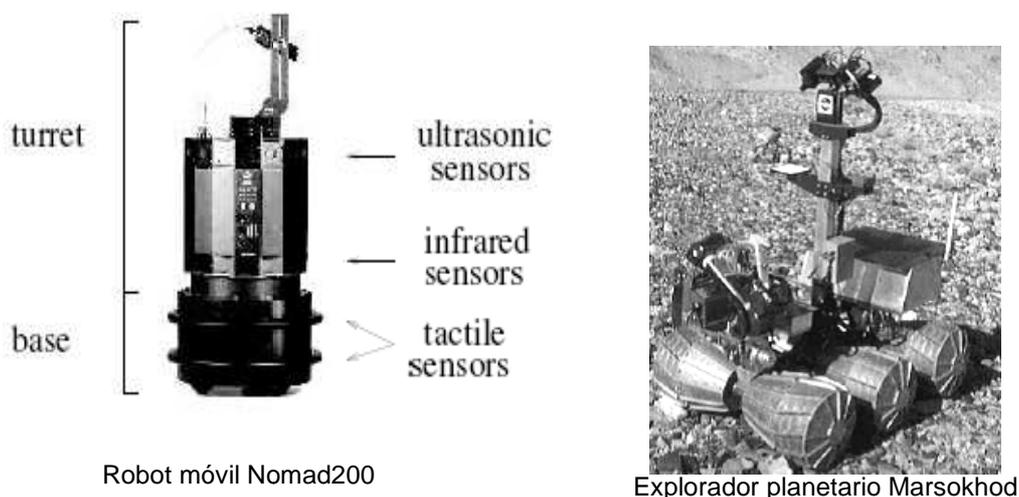
**4.2.2. Selección Estructura  $\lambda$  de HMM para Robot Reconocedor de lugares.** Al igual que el proceso de reconocimiento de voz visto anteriormente, definir la estructura del HMM para el aprendizaje y reconocimiento de lugares con robot móviles, requiere de establecer que es lo que un HMM representa dentro de la aplicación y cómo lo hace. Algunos pasos para definir la estructura son los siguientes:

Definir el número de características del terreno a reconocer y la forma de describirlas. Las características son seleccionadas por el hecho de que sean repetibles y observables por el ser humano para propósitos de validación, evaluación, entrenamiento y etiqueta.

Encontrar el modelo más apropiado HMM para representar cada una de las características identificadas en el punto anterior. Esto implica seleccionar la topología, el número de estados, la dimensión de las observaciones (número de sensores a evaluar), y su distribución en cada uno de los estados.

En este documento se describirá los casos generales de aprendizaje y reconocimiento de características tomado del trabajo de Aycard [2] para ambientes estructurados en interiores (*indoor environment*) y de ambientes en exteriores (*outdoor environment*) con HMM de segundo orden (HMM-2). Sin embargo, en este documento se hará mención a las HMM's de primer orden que no dista mucho en estructura y algoritmo del HMM-2. Los robots utilizados por tal autor para esos ambientes se describen brevemente y se ilustran en la Figura 31; En el reconocimiento de ambientes interiores utilizó un Nomad200 compuesto por una base consistente en 3 ruedas y sensores táctiles y una torre poligonal de 16 lados, giratoria sobre la base y compuesto de sensores infrarrojos y ultrasonido en cada uno de sus lados. En el reconocimiento de ambientes exteriores utilizó el explorador planetario, *The Marsokhod rover*. compuesto de 6 ruedas de movimiento independiente, con tres segmentos de chasis articulatorias, sensores de corriente para cada motor de la rueda, un sensor de balance (roll sensor) que mide el ángulo entre el eje vertical del móvil y el eje perpendicular a la dirección de navegación, y un sensor de nivel del frente del móvil (pitch sensor).

Figura 31. Robots móviles para reconocimiento de lugares interiores y exteriores<sup>25</sup>.



Se define a continuación los pasos para definir la estructura de HMM.

**4.2.2.1. Definición del Conjunto de Características del Terreno.** La definición del conjunto de características de un ambiente debe proporcionar una completa descripción de lo que el robot móvil puede ver durante un recorrido. Su selección está en función del conocimiento que pueda tener el diseñador acerca del terreno, definiendo reglas fundamentales para identificar cada característica en base a las variaciones de los sensores que lo constituyen. Estas reglas son solo para uso humano en etapas de segmentación y etiquetado de las tramas de datos del cuerpo de entrenamiento. Si existen características que no son reconocidas por el robot, estas son tratadas como ruido. Después de definir los elementos del conjunto de características, cada uno de estos será representado por los parámetros de un solo HMM específico. Por lo tanto habrá tantos HMM's como características se definan.

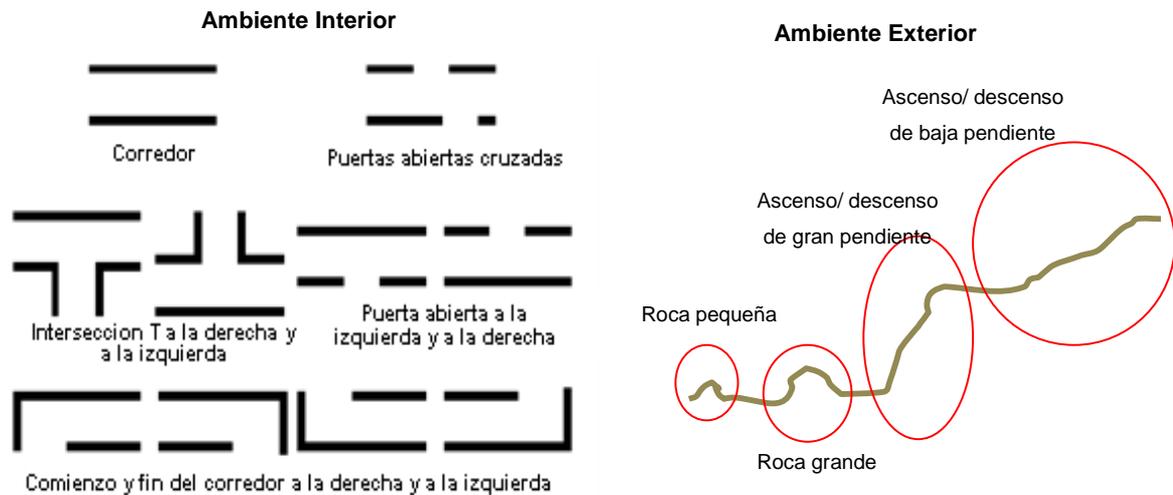
Las características detectadas de un terreno son muy variables y pueden ser definidos según varios criterios. Uno de ellos es la detección de características **naturales o**

---

<sup>25</sup> Tomado de: Aycard, Olivier; Mari, Jean- Francois; Washington, Richard. *Learning to automatically detect features for mobile robots using second-order Hidden Markov Models*. En: International Journal of Advanced Robotic Systems. Vol. 1, No. 4. (2004). p. 239 y 243.

**artificiales**, considerándose a las primeras como aquellas que ya existen dentro del ambiente y la ultimas como aquellos que son adicionados por el diseñador para su fácil detección. Sin embargo, las características artificiales son inútiles en terrenos desconocidos y remotos debido a la imposibilidad del diseñador de adicionarlos a estos. Otro criterio es la de definición de características del terreno en forma **manual o automática**, en donde el conjunto de características pueden ser definidas en la maquina usando la percepción de los humanos sobre las condiciones del terreno, o por el contrario, las maquinas deben descubrir las diferentes situaciones automáticamente con escasa participación del diseñador. La definición de característica en forma automática por parte del robot móvil, hace que además de utilizar un HMM para su reconocimiento, se acuda a alguna técnica adicional de aprendizaje no supervisado para la identificación de las clases que componen el terreno en forma análoga al reconocimiento de voz para la identificación de clases de sonidos. Un último criterio, es la de considerar que muchas veces las características del terrenos se identifica por medio de una **secuencia temporal** de la información del sensor, y no simplemente un dato independiente o una captura instantánea. Por lo general, las aproximaciones con secuencias temporales son más robustas que las detecciones instantáneas por su capacidad de representar la evolución de las características detectadas pasadas y descubrir su correlación con aquellas detectadas actualmente. En este criterio, se basa la fortaleza de aplicar una Cadena Oculta de Markov.

En la Figura 32 se observa las diferentes características definidas para ambientes interiores y exteriores. Para ambientes interiores se observa características comunes al que se denominan “lugares” por los cuales el robot se enfrentará. Ejemplo de lugares se encuentran los corredores, intersecciones, esquinas iniciales y finales, puertas abiertas, etc. En ambientes exteriores se tendrán características llamadas “situaciones” que difieren al caso de ambientes internos, enfrentándose por ejemplo a cruce de rocas grandes y pequeñas, ascensos y descensos de lomas de diferentes pendientes, etc.



**Figura 32. Set de características del terreno en ambientes interiores y exteriores.**

Para describir cada lugar o situación, se necesita obtener la evolución de las medidas conjuntas de los sensores seleccionados cuando navega mientras detecta cada característica. Esta es una tarea compleja porque las medidas de los sensores para ambas situaciones son ruidosas y por que al mismo tiempo que hay un lugar o situación en el lado derecho o izquierdo del robot, hay otra situación distinta en el lado opuesto.

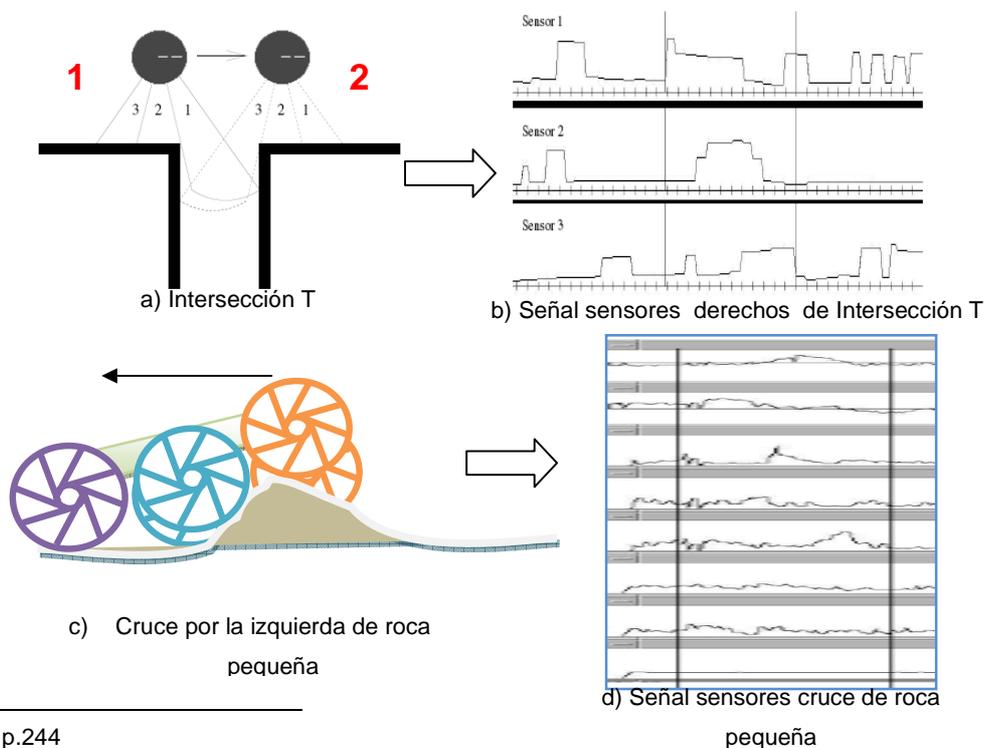
Aycard<sup>26</sup> describe una característica particular como una intersección T para un ambiente interior como la que se observa en la Figura 33.a. Se detalla que en el lado derecho del robot se utilizan 3 sensores ultrasonido que detectan el lugar cuyas medidas se observa en la Figura 33.b, revelando cambios durante el recorrido de dicha característica del punto 1 al punto 2. Las reglas asumidas para identificar este lugar basado en las medidas de los sensores se pueden resumir así: *Se detecta el comienzo de una intersección cuando la medida del primer sensor se incrementa de forma repentina, y el final de dicha intersección cuando la medida del tercer sensor decrece repentinamente. Mientras ocurren estos dos acontecimientos, la medida del segundo sensor presenta un incremento y un descenso.* De manera similar, se debe realizar el mismo procedimiento para obtener reglas en otros lugares diferentes a la intersección,

<sup>26</sup> Ibid., p.240

analizando la información de los sensores para identificar comportamientos particulares en sus medidas.

Por otro lado, Aycard<sup>27</sup> también describe una situación para ambientes exteriores como lo es cruzar una roca pequeña en uno de sus lados. En la Figura 33.c y Figura 33.d se observa esta situación con las señales obtenidas de 8 sensores (Roll, pitch y las corrientes de las 6 ruedas) cuando esto ocurre. En base a estas señales, se estipula las reglas que generalizaría dicho evento de la siguiente manera: *Cuando un robot cruza una roca pequeña en cualquiera de sus dos lados, se distingue que el sensor de balance (roll sensor) muestra un suave incremento cuando comienza a escalar la roca, luego ocurre un pequeño y rápido descenso cuando abandona la roca. Estas variaciones igualmente ocurren secuencialmente en las corriente de la rueda frontal, media y trasera. El sensor de nivel frontal (pitch sensor) siempre muestra un pequeño incremento, luego un pequeño y rápido descenso y finalmente un pequeño incremento.* Igualmente con las otras situaciones se estipulan reglas para su identificación en base a las señales de los sensores.

**Figura 33. Evolución temporal de características y Medidas de los sensores.**



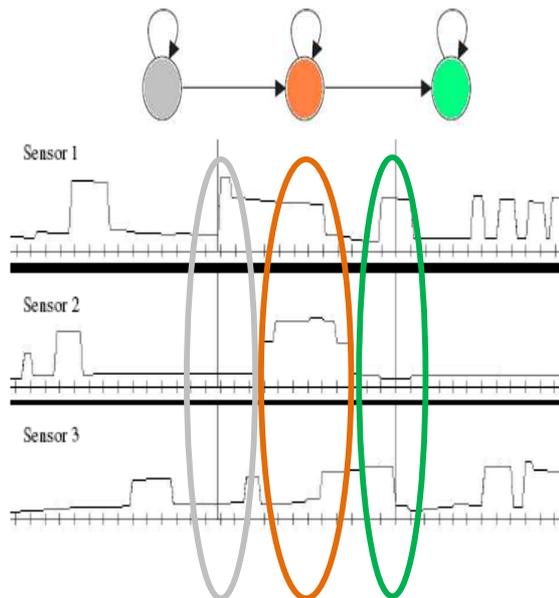
<sup>27</sup> Ibid., p.244

Es de notar que la estipulación de las reglas apropiadas para la identificación de las diferentes características requiere en gran parte del hábil trabajo del diseñador en base a su conocimiento de los eventos ocurridos dentro del terreno. Este trabajo es necesario tanto para la segmentación y entrenamiento del HMM en una característica del terreno específica como para la validación y evaluación de la eficiencia del reconocimiento del robot a dicha característica.

En comparación con el reconocimiento de voz, la tarea de definición de características del terreno para un robot móvil equivale a identificar el tamaño del vocabulario a ser utilizado, entendiéndose como todas las posibles eventos (número de HMM's) que se podrían presentarse dentro de la aplicación a desarrollar. Por lo tanto, después de identificados el dominio de eventos en esta aplicación en particular, es necesario definir la topología, número de estados, observaciones y su distribución de probabilidad.

**4.2.2.2. Selección del número de Estados y topología.** En muchas aplicaciones con HMM's se ha mostrado la efectividad de utilizar la topología de Izquierda-derecha, pues esta describe los cambios naturales de un evento como un conjunto de sub-eventos discretizados, únicos e irrepetibles después de que estos hayan culminado dentro del mismo evento. En el reconocimiento de lugares o situaciones del terreno, la descripción temporal de los datos de cada una de las características hace notar que después de ocurridos cierta combinación de cambios en las medidas de los sensores, es improbable que estos cambios vuelvan a ocurrir durante la evolución de una característica después de finalizados dichos cambios. En términos generales, en la topología de izquierda- derecha, un estado dentro del HMM no podrá repetirse luego de que se haya abandonado dicho estado para continuar con uno nuevo. Retomando la regla estipulada anteriormente para el reconocimiento de una intersección, no tendría sentido por ejemplo, que el robot reconozca que se encuentra en medio de una intersección y que posteriormente se encuentre en los sensores información del inicio de dicha intersección sin antes haber llegado a reconocer el final de la intersección

La selección del número adecuado de estados que conforman la descripción temporal de la característica del terreno a reconocer, no es un procedimiento directo y determinista, sino que por el contrario, la iteración, el tanteo y error se aplican para analizar los posibles cambios (o sub-eventos) temporales y significativos que podrían conformar dicha descripción que proporcione una tasa de reconocimiento óptimo. Un primer paso para iniciar estipulando el número de estados es analizar el conjunto de señales obtenida de los sensores junto con las reglas estipuladas para cada característica del terreno. En el caso del reconocimiento de la intersección T, se podría pensar intuitivamente en la existencia de 3 estados en el cual el primero de ellos contiene los incrementos del primer sensor correspondiente al inicio de dicho lugar, luego el segundo estado contendría la parte estacionaria de las señales (en el medio de la intersección) y finalmente un tercer estado que contiene el final del lugar representado por la señal decreciente del tercer sensor. Posteriormente se recurre a la adición o eliminación de estados dependiendo de la eficiencia en el reconocimiento de dicho lugar. En la Figura 34 se describe la correspondencia del número de estados y la topología seleccionada para un lugar en reconocimiento.



**Figura 34. Topología de estados y evolución temporal de las señales de sensores.**

El número de estados definidos para en el reconocimiento de ambientes exteriores es un poco más complejo de analizar. En forma recurrente, se podría pensar en analizar los diferentes sub-eventos que se presentan en el reconocimiento de una situación al igual que el reconocimiento de la intersección, teniendo en cuenta las señales de los sensores. Se encuentra en este caso que el número de sensores, a comparación del robot móvil para interiores, es mayor y por lo tanto, el número de sub-eventos importantes está relacionada a las posibles combinaciones de cambios importantes durante la evolución del reconocimiento de tal situación. Se determina entonces, que entre más sensores tenga un robot, el número de estados tiende a ser mayor por la ampliación del rango de posibles eventos. Se estipula, por lo tanto que el mejor número de estados para el HMM del explorador planetario es de 5, pues es este el que mejor proporciona eficiencia en el reconocimiento<sup>28</sup>.

**4.2.2.3. Definición del conjunto de observaciones y su distribución de probabilidad.** La dimensión del vector de observaciones depende del número de sensores que se utilicen en cada uno de los robots. Para el caso del Nomad200 el HMM se utiliza información de 16 sensores ultrasónicos, y para el Marsokhod se utiliza información de 1 sensor de nivel frontal (pitch), 1 sensor de balance (roll), y 6 medidas de corrientes correspondiente a cada una de las ruedas.

Es evidente además, que el uso de funciones de distribución de probabilidad gaussianas para variables continuas es mandatorio por la naturaleza no discreta de las señales obtenidas por los sensores. En la Figura 35 se representa una manera de ver las distribuciones de probabilidad por estado para observaciones multidimensionales (6 sensores) en la descripción de una característica de un ambiente interior como lo es una intersección T.

---

<sup>28</sup> Ibid.,p.245

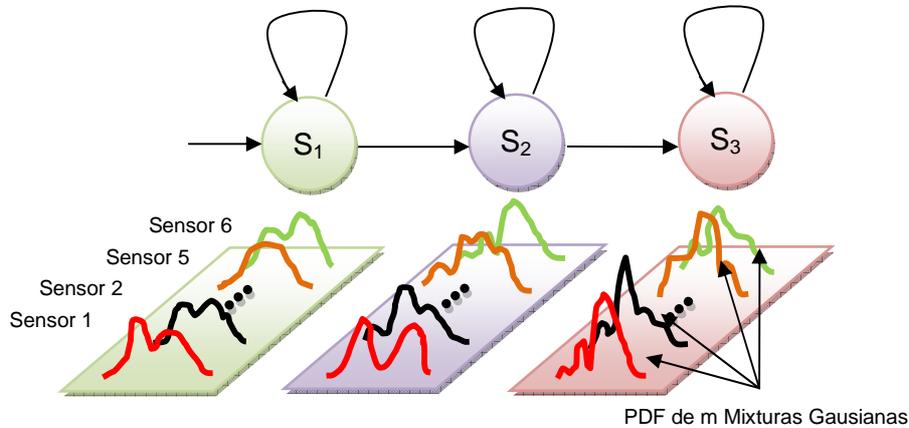


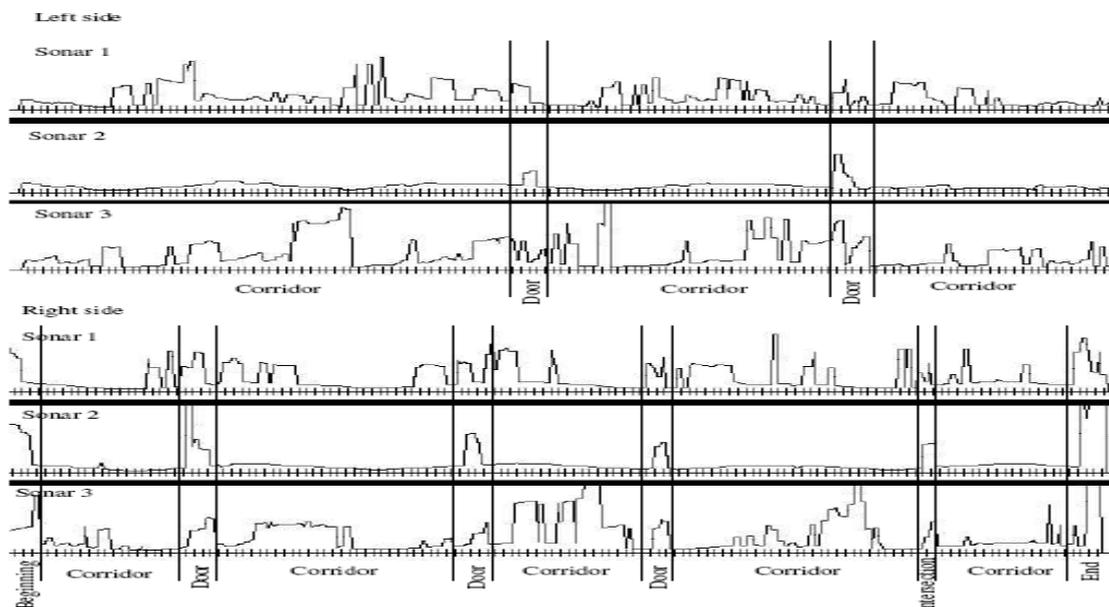
Figura 35. Funciones de distribución de probabilidad por estado Intersección T.

**4.2.3. Entrenamiento de los modelos y Etapa de Reconocimiento.** Se debe construir un cuerpo de entrenamiento para el modelo HMM que represente cada uno de los lugares o situaciones. Para esto, el robot móvil debe hacer varios recorridos en espacios que contengan combinaciones de todas las situaciones y lugares (ver Figura 36), según sea el caso, y almacenar todas las secuencias de datos obtenidas en los sensores. Después de dicha recolección, se hace la segmentación y etiqueta de cada lugar o situación correspondiente a partir de las reglas estipuladas anteriormente para cada uno de ellos.



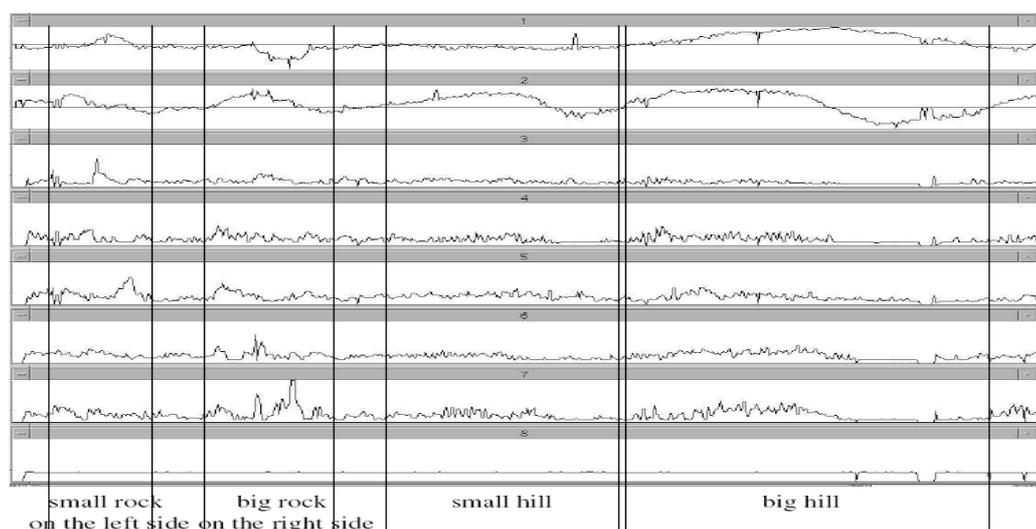
Figura 36. Secuencia combinada de varios lugares en ambientes interiores.

Por último, se entrena cada modelo HMM con el respectivo conjunto de datos segmentados y etiquetados para cada característica. En la Figura 37 se ilustra el proceso de segmentación de las diferentes características a través de las señales de los sensores para ambientes interiores. Se observa que al lado izquierdo del robot se detecta la secuencia Corredor-Puerta-Corredor-Puerta-Corredor y al lado derecho la secuencia Comienzo-Corredor-Puerta-Corredor-Puerta-Corredor-Puerta-Corredor-Intersección-Corredor- Fin.



**Figura 37. Segmentación y Etiqueta de los lugares en ambientes interiores.**

En la Figura 38 se observa el proceso de segmentación y etiqueta de situaciones en ambientes exteriores con las señales de los 8 sensores (ordenadas descendientemente así: Roll, Pitch, 3 corrientes ruedas izquierdas, 3 corrientes ruedas derechas), cuya secuencia de situaciones se identifica: Roca pequeña en el lado izquierdo- Situación normal- Roca grande en el lado derecho- Situación normal- pequeña pendiente- Situación normal – gran pendiente.



**Figura 38. Segmentación y Etiqueta de Situaciones en Ambientes Exteriores.**

Luego de la etapa de entrenamiento, se aplica la etapa de reconocimiento donde se busca no solamente identificar característica en forma individual, sino identificar un lugar diverso en el cual combine una serie de características del terreno.

La equivalencia de esta situación se puede asociar al reconocimiento de oraciones pronunciadas por un locutor, en donde la interconexión seriada de cada uno de los modelos HMM que representan una palabra específica, está gobernada por reglas gramaticales que componen tal oración. Con esta comparación, el ambiente el cual enfrenta el robot se describe por medio de reglas “gramaticales” que restringen el conjunto de secuencias posibles de modelos HMM que lo componen tal ambiente específico. Usando esta gramática, todos los modelos HMM de cada característica se unen a un modelo HMM superior que los contenga, descifrando la secuencia de lugares o situaciones por medio del algoritmo de Viterbi.

Más formalmente, esta gramática representa todas las posibles cadenas de Markov correspondientes a la parte oculta de los modelos conectados. En este caso cada nodo

corresponde a HMM's asociados con una característica en particular. Límites entre dos HMM's corresponde una unión entre el ultimo estado de uno de los HMM's y el primerio estado del siguiente modelos HMM. La probabilidad asociada a cada uno de los límites representa la probabilidad de que exista la unión entre ellos (transición entre HMM's). Luego, la secuencia de estados (HMM's) más probable, determinada por el algoritmo Viterbi, estipula la lista ordenada de características que el robot vio durante su recorrido. Se debe notar que la lista de modelos es conocida solo cuando el recorrido es completado. Se hace la hipótesis que 2 o más modelos de las características no se pueden solaparse. El uso de la gramática en esta aplicación tiene una ventaja: permite la eliminación de algunas secuencias que nunca ocurrirán en el ambiente, y desde el punto de vista computacional, la gramática evitara algunos cálculos inútiles.

La gramática puede ser dada en forma previa o aprendida. Para aprender la gramática, se usa los primeros modelos y se estiman en datos no segmentados como en la etapa de reconocimiento. Específicamente, se conectan todos los modelos vistos por el robot durante un recorrido completo a través un modelo HMM mayor correspondiente a la secuencia de elementos observados y se entrena el modelo resultante con los datos no segmentados.

En la Figura 39 se ilustra el reconocimiento de una situación real compuesta por varias características definidas para un ambiente interior, cuya identificación está regida por la gramática de tal situación modelada por un HMM de mayor jerarquía. El resultado es la selección de la secuencia de características más probables arrojadas por el algoritmo de Viterbi.

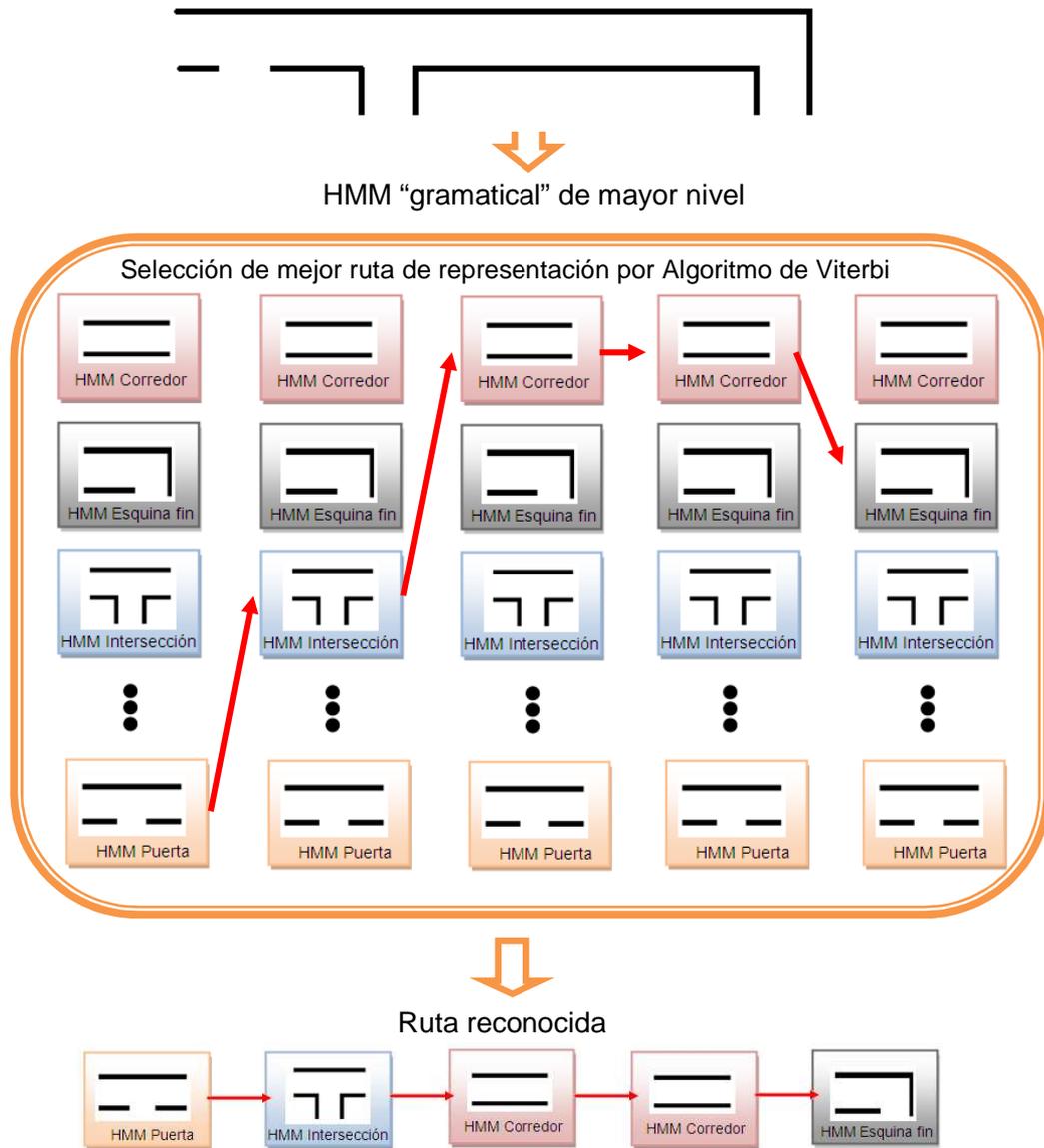


Figura 39. Reconocimiento lugares compuestos con algoritmo de Viterbi.

### RESUMEN APRENDIZAJE Y RECONOCIMIENTO DE LUGARES CON ROBOTS MÓVILES CON HMM.

#### 1. Naturaleza y dificultades del reconocimiento de lugares con robots móviles.

Permitir a un robot tener total autonomía para navegar por lugares de diferentes condiciones en ambientes interiores y exteriores a partir de los sensores que posee. La dificultad para interpretar la información de los sensores se fundamenta en:

- Señales ruidosas o incompletas detectadas en los sensores.
- Confusión perceptual de lo que observa.
- Excesiva cantidad de datos en tiempo real y pobre correlación entre ellos.

#### 2. Estructura de HMM.

- **Lo que representa un HMM:** Dependiendo de la cantidad de situaciones o lugares (pequeñas etiquetas a sitios) que se represente, se puede designar a un HMM a cada una de estas situaciones describiendo la evolución temporal de la señal de los sensores a medida que avanza por el terreno.
- **Topología:** Por lo general se utiliza la de Izquierda - a - Derecha (Left- to Right) por representación sencilla del flujo evolutivo como secuencia temporal de medidas sensoriales.
- **Numero de estados:** Dependiendo de los cambios importantes de las señales en cada uno de los HMM seleccionados, se designa el número de estados en base a reglas estipuladas previamente para su identificación o al tanteo hasta que se obtenga un número adecuado que mejor probabilidad represente. Se utilizan algunas veces técnicas de aprendizaje no supervisado, donde la identificación de clases previas para el entrenamiento lo hace automáticamente el robot.
- **Observaciones O:** equivale a las variaciones de los sensores y su dimensión depende del número de sensores que trabajaran para el reconocimiento. La distribución de probabilidad de las observaciones dentro de los estados está compuesta por funciones gaussianas de múltiples mezclas para cada una de las variables que conforma la dimensión del vector de observaciones de medidas sensoriales.

#### 3. Entrenamiento de los Modelos HMM

El robot móvil debe hacer varios recorridos en espacios que contengan combinaciones de todas las situaciones y lugares. Luego se grafica todas las señales de los sensores, se segmentan los lugares y situaciones importantes en base a las reglas estipuladas, se etiqueta y se almacena toda la información para una misma situación o lugar y por último se entrena cada uno de los HMM's específicos con sus datos segmentados correspondientes.

#### 4. Reconocimiento de Lugares con robots móviles.

El mismo procedimiento de entrenamiento es aplicado en el recorrido de combinación de situaciones (HMM's seriadas) en donde la conexión entre HMM's para describir un lugar real es calculada por el algoritmo de Viterbi seleccionando la ruta más probable con respecto a un lugar real en reconocimiento.

### 4.3. TÉCNICAS PREDICTIVAS DE ACTIVIDADES DE MANTENIMIENTO EN LA INDUSTRIA: DIAGNOSTICO Y PRONOSTICO CON HMM'S

Todo equipo mecánico, eléctrico o electrónico que se encuentre instalado dentro de algún proceso industrial esta propenso a sufrir fallas que atenten contra el estado de mismo y la función que desempeña dentro del proceso productivo. Tales fallas, en consecuencia, afectan la continuidad productiva de todo el sistema, producen pérdidas sobre los recursos en términos de dinero y tiempo, y además ponen en riesgo la seguridad del recurso humano cercano al proceso.

Por tal motivo, está el interés de la comunidad científica investigar sobre métodos que apoyen a la predicción y prevención de fallas de equipos, sobre todo de aquellos que manifiesten alta criticidad dentro del proceso y que afecten con más fuerza a las variables anteriormente mencionadas. Algunos casos de alta criticidad se pueden evidenciar en el sector eléctrico en donde activos tales como transformadores de potencia, interruptores, seccionadores, entre otros, se deben mantener de tal forma que aseguren la continuidad y disponibilidad del flujo eléctrico, evitando colapsos sobre el sistema interconectado. Se presenta igualmente, alta criticidad en la industria de la aviación, en donde las aeronaves deben mantenerse en óptimas condiciones para asegurar la integridad y vida de los pasajeros durante el vuelo. En los centros médicos, los equipos que apoyan en funciones vitales a los pacientes no deben dar oportunidad a ningún tipo de fallo que atente contra sus vidas, entre otros casos.

Actualmente, está tomando fuerza como método de prevención de fallas inminentes, el mantenimiento basado en condición (CBM) o Mantenimiento predictivo (PdM) en el cual tienen en cuenta el estado de deterioro actual del equipo.

En las siguientes secciones se mencionaran aspectos sobre algunas filosofías de mantenimiento que conllevan a pensar en técnicas más inteligentes de análisis de condición de equipos, se mencionará la importancia del mantenimiento predictivo o CBM junto con sus fundamentos, componentes y modos de asumir las fallas. Posteriormente, se mostrará la aplicación de Cadenas Ocultas de Markov en las etapas de Diagnostico y Pronostico de un CBM para condiciones de operación de un

sistema de transmisión de un helicóptero Westland, la identificación de la estructura del HMM (sus estados, topología, observaciones y FDP's), sus pruebas de entrenamiento, y por último, la evaluación de Diagnostico y Pronostico para diferentes condiciones de operación del helicóptero.

**4.3.1. Filosofía de mantenimiento de equipos industriales.** Las actividades de mantenimiento a equipos se han estado practicando desde la antigüedad, cuando los egipcios y los romanos hacían mantenimiento a sus infraestructuras de acueductos, pirámides, arsenales, etc., para preservar la integridad de aquellos. Es hasta la época de la II Guerra Mundial que se comenzó a desarrollar el concepto de mantenimiento dado que las consecuencias del conflicto armado incrementó la necesidad de productos de toda clase mientras que la mano de obra industrial bajaba drásticamente. Este panorama obligó a los industriales a crear sistemas eléctricos y mecánicos automatizados complejos que sustituyeran la ausente mano de obra y suplir los necesitados productos.

A partir de ese momento las industrias habían comenzado a depender de aquellos equipos y el daño o deterioro de alguno de ellos representaba tiempo improductivo y consecuente pérdida de dinero. Ante estas eventualidades, se procedía a corregir la falla con la reparación o sustitución del equipo corriendo con los costos necesarios para su ejecución. Este era una práctica muy básica de mantener un equipo en servicio y es conocido hoy día como **mantenimiento correctivo** (CM: Corrective Maintenance), donde se espera simplemente que el equipo falle para repararlo.

En forma intuitiva se llegó a la idea que las fallas se podían y se debían prevenir, creando el concepto de **mantenimiento preventivo** (PM: Preventive Maintenance)<sup>29</sup>. El objetivo de este concepto era estipular programas periódicos que sacaban de servicio por un pequeño lapso de tiempo el equipo correspondiente y permitir hacer una revisión general o sustitución de algunos componentes sin importar el estado actual de deterioro es que se encontraran. Sin embargo, detener el proceso productivo

---

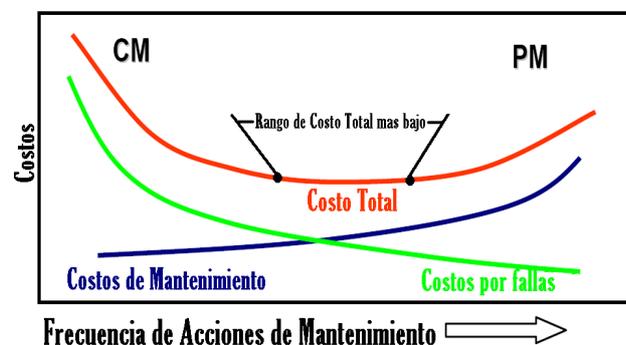
<sup>29</sup> UNIVERSITAS. *Mantenimiento su implementación y gestión*. Argentina: Universitas, marzo,2005. p.121-136.

muy frecuentemente y sustituir componentes en forma periódica sin conocer su estado real de deterioro significaban costos adicionales al proceso de producción.

Así pues, estos dos conceptos han sido los límites para tomar una decisión entre no hacer revisiones preventivas y esperar a que se deteriore el equipo para repararlo, o hacer “algo” antes de que ocurra la falla con una frecuencia de ejecución que debe ser elegida apropiadamente. Sea cual sea la decisión, la técnica de mantenimiento seleccionada agregará un costo adicional al proceso y el reto de las empresas es averiguar cuál sería aquella que menos costos y riesgos le genere. Entonces la pregunta que siempre busca resolverse para la planeación previas de recursos en el mantenimiento de un activo es: **¿Cuándo ejecutar las tareas de mantenimiento de un equipo determinado obteniendo el menor costo- riesgo?**

En la Figura 40 se ilustra el esquema mencionado por el cual se obtiene el costo total de mantenimiento<sup>30</sup>. Se observa que la baja frecuencia de actividades de mantenimiento (con tendencia a CM) representa un alto costo en reparación y pérdidas productivas causados por las fallas, pero un bajo costo en uso recursos de mantenimiento. Por otro lado, una frecuencia alta de actividades (aumento de PM) representa un bajo costo por consecuencia de fallas pero un incremento de costo por recursos de mantenimiento y horas de indisponibilidad del equipo. Está claro entonces que la tarea esta en seleccionar un rango óptimo de frecuencia de acciones de mantenimiento cuyo costo total sea bajo y se adapte a las condiciones reales del proceso, asegurando la integridad física de los humanos.

**Figura 40. Selección de costos óptimos en mantenimiento de equipos**

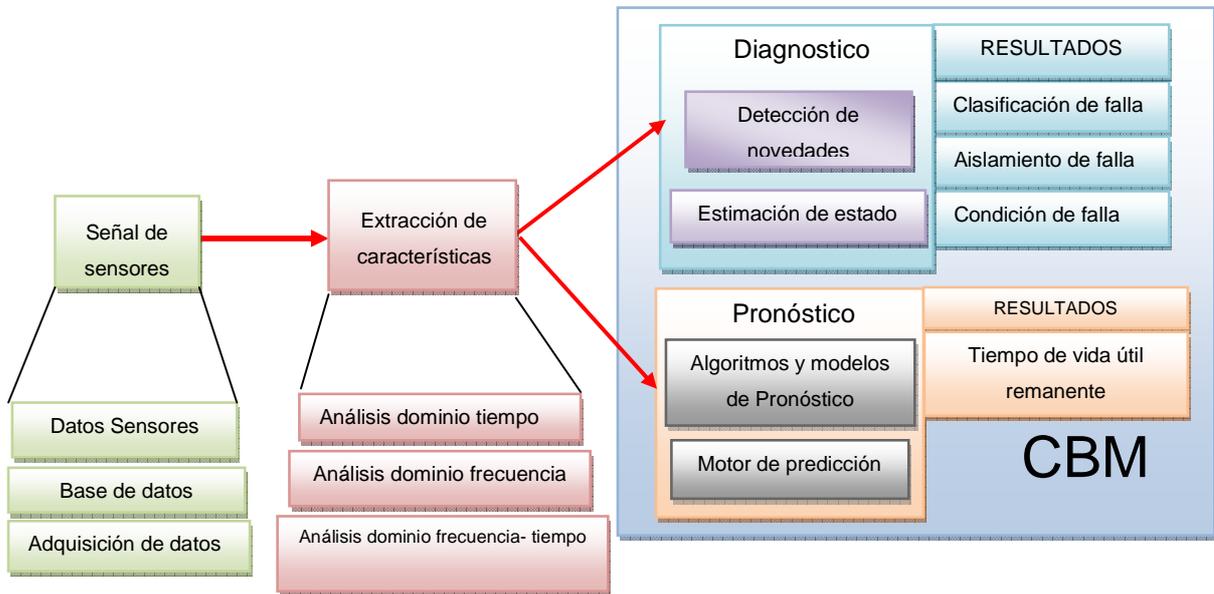


<sup>30</sup> LEVITT, Joel. Preventive and Predictive Maintenance. New York: Industrial Press, 2005. p.12

Con el fin de hacer más fácil esta tarea de selección, se han estado desarrollando técnicas de mantenimiento para el monitoreo de condiciones de equipos, que a diferencia del mantenimiento preventivo convencional, tienen en cuenta el estado actual de deterioro del equipo para decidir en qué momentos es necesario hacerle mantenimiento. En este caso no se estipula periodos fijos de mantenimiento sino que en base a la condición actual detectada del equipo por medio de información captada por sensores, equipos de medición especializados, base de datos, etc., se toma la decisión acerca del momento idóneo en el cual una actividad de mantenimiento podría ser ejecutado. Esta técnica es conocida como Mantenimiento Basado en Condición (CBM: Condition Based Maintenance) o Mantenimiento Predictivo (PdM: Predictive Maintenance).

La implementación de este método está fundamentado en la utilización de tecnología sofisticada para monitoreo, haciendo que se enfoquen generalmente solo a equipos de alta criticidad por su costo de implantación. Algunos de los equipos de monitoreo de condiciones realizan mediciones de variables como temperatura a través de termografía, vibraciones por medio de acelerómetros y medidor de espectros, fugas a través de sensores ultrasonido, análisis de gases disueltos en aceites de aislación eléctrica en transformadores o interruptores, entre otros.

Las mediciones obtenidas por los sensores para el monitoreo de condiciones de equipo son utilizados para fines de Diagnostico de fallas y Pronostico de eventos. El proceso de **Diagnostico** detecta la presencia de falla y hace su posterior identificación dentro de un sistema (FDI: Fault Detection and Identification). El proceso de **Pronóstico** indica la predicción del tiempo faltante de vida útil del equipo a partir de las condiciones actuales del equipo. En la Figura 41 se observa los elementos básicos utilizados en el Mantenimiento Basado en Condición.



**Figura 41. Estructura de Mantenimiento Basado en Condición.**

Sus etapas básicamente consisten en la adquisición de datos a través de sensores, bases de datos, etc., sobre las variables del equipo, que a su vez pasan a ser procesados a través de una etapa de extracción de características con la utilización de métodos de análisis en el dominio del tiempo, en dominio de la frecuencia, o la combinación de ambas. Posteriormente, se aplican las características a algoritmos de Diagnostico y Pronóstico para la detección e identificación de fallas y el cálculo de tiempo de vida útil restante del equipo, respectivamente. El resultado final obtenido de todo este proceso es el listado de todas las acciones recomendadas para la ejecución del mantenimiento del equipo monitoreado. En este documento solo se hará énfasis en las etapas de Diagnostico y pronóstico en el análisis de las condiciones de operación de equipos.

La esencia del CBM consiste en suponer que ningún equipo falla en forma repentina, sino que por el contrario, presenta una evolución de síntomas de deterioro que se manifiestan en el tiempo y que en etapas tempranas algunas de estas no son detectadas a simple vista<sup>31</sup>. Por lo general, los operadores de los equipos detectan

<sup>31</sup> ALLIED RELIABILITY. *PdM Secret Revealed: How to improve you PdM program Or Start Ones from Scratch*. s.L: s.n, march 28,2006. p.4-5.



### 4.3.2. Selección de la Estructura de un HMM para diagnostico y pronostico.

Como caso real de aplicación se tomará como ejemplo el documento “*Condition based maintenance of machine using Hidden Markov Models*”, de los autores Carey Bunks y Dan MacCarthy<sup>32</sup> en cual aplican Cadenas Ocultas de Markov para el diagnostico y pronostico de condiciones de operación en una caja de transmisión de un helicóptero del fabricante Westland. Algunos modelos de este fabricante se observan en la Figura 43.



**Figura 43. Modelos de Helicopteros fabricante Westland<sup>33</sup>.**

El método utilizado para llevar a cabo el CBM en la caja de transmisión del Westland se apoya en el uso de medidas de vibración obtenidas de 8 sensores acelerómetros ubicados en diferentes posiciones del sistema de transmisión como se muestra en la columna izquierda de la tabla 1. Estos están instalados de tal forma que detecten 68 diferentes condiciones de operación los cuales están constituidos por 9 niveles de torque, 8 diferentes defectos y dos niveles de severidad para cada defecto. Los niveles de torque son valores porcentuales de torque nominal y se utilizan el 27, 40, 45, 50, 60, 70 75, 80 o 100%. Los 9 tipos de defectos y sus índices se muestran en la tabla 1 (columna derecha).

| <b>POSICIÓN DE SENSORES</b>     | <b>DESCRIPCIÓN DE DEFECTOS</b>                       |
|---------------------------------|--|
| 0 Tacómetro                     | 1 Sin defectos (piezas originales).                  |
| 1 Entrada del motor de estribor | 2 Corrosión del rodamiento del engranaje planetario. |

<sup>32</sup> BUNKS, Carey et. al. Condition-Based Maintenance of Machines using Hidden Markov Models. En: Mechanical Systems and Signal Processing. (march, 2000); p.597-612.

<sup>33</sup> Imágenes tomadas de: [www.wikipedia.org](http://www.wikipedia.org). Palabra clave: Westland

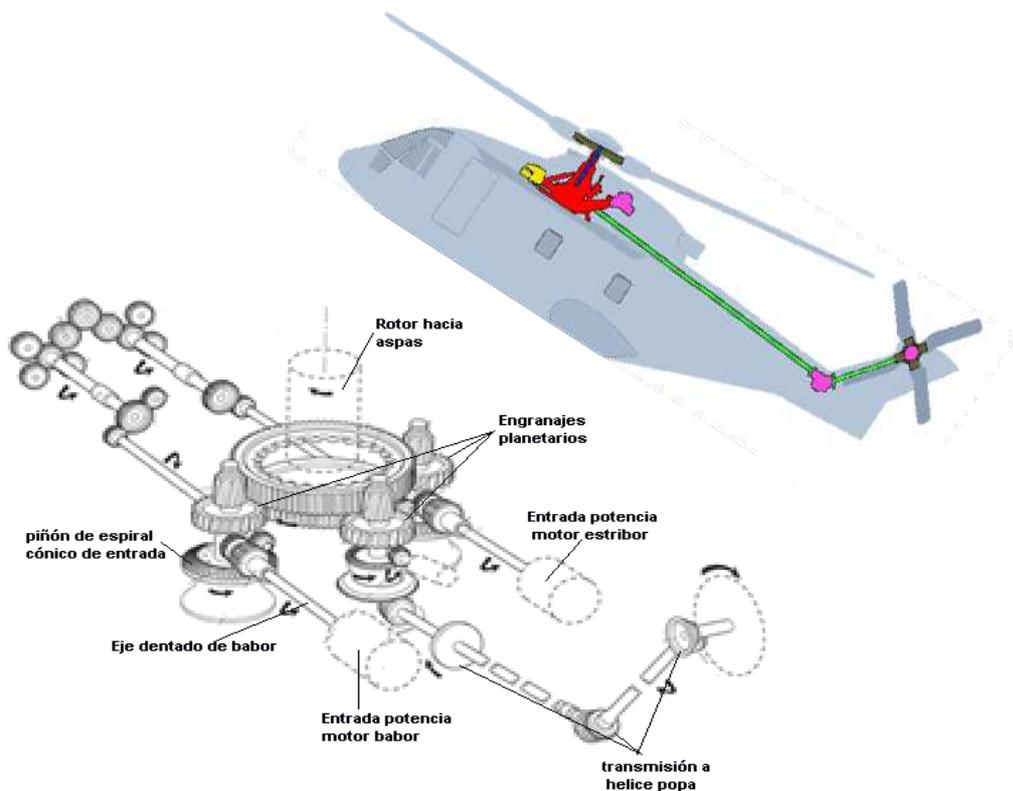
## Hidden Markov Models (HMM's) y aplicaciones

|   |                            |   |   |
|---|----------------------------|---|---|
| 2 | Entrada del motor de babor | 3 | Corrosión del rodamiento del piñón de entrada.            |
| 3 | Caja mixta de popa         | 4 | Desgaste del piñón de espiral cónico de entrada .         |
| 4 | Eje dentado de estribor    | 5 | Rotura o mella en el piñón helicoidal de entrada.         |
| 5 | Planetario de estribor     | 6 | Propagación de fractura del engranaje helicoidal inactivo |
| 6 | Planetario de babor        | 7 | Propagación de fractura de la transmisión colectora       |
| 7 | Eje dentado de babor       | 8 | Propagación de fractura del eje dentado                   |
| 8 | Controlador de accesorio   | 9 | Sin defectos (después de reparación)                      |

Tabla 1. Descripción de defectos y posición de sensores.

Un esquema generalizado del sistema de transmisión de un helicóptero Westland se observa en la Figura 44.

**Figura 44. Sistema de transmisión de un Helicóptero Westland<sup>34</sup>.**



<sup>34</sup> Imágenes tomadas de: [www.westlandhelicopters.com](http://www.westlandhelicopters.com)

Para cada uno de los defectos vistos en la tabla1 se evalúan 2 niveles de severidad que hacen referencia al grado de deterioro por el que se desarrolla temporalmente tal defecto, infiriendo por lo tanto, la no existencia de una secuencia de severidad del nivel 2 (Level II) que sería el más grave, al nivel de severidad 1 (Level I), pues no es realmente posible que una pieza se regenere, a menos que se le haga una sustitución o acondicionamiento. Cabe mencionar que ante la ausencia de defectos, no se presenta ningún nivel de severidad.

**4.3.2.1. Naturaleza de los datos de vibración del sistema de transmisión.** Para determinar que las variables seleccionadas de torque, tipos de defectos, y niveles de severidad son suficientes para detectar la condición actual del sistema y la estimación de la vida restante del mismo, los autores analizan tres tipos de situaciones que demuestran la posibilidad de aplicar Cadenas de Markov para el diagnóstico y pronóstico del sistema de transmisión. Para esto, se analizan los espectros de las señales de los sensores en diferentes condiciones de operación de torque, defectos y nivel de severidad. Con estas pruebas se demuestran la propiedad cuasi-estacionaria de la señales de una condición de operación ante diferentes maniobras del helicóptero; el funcionamiento de la tarea de diagnóstico al detectar diferencias entre situaciones sin defectos y con defectos del sistema de transmisión; y el funcionamiento de la tarea de pronóstico al detectar la evolución de los defectos en niveles de severidad y predecir el tiempo de vida útil restante. Las situaciones son las siguientes:

**Situación 1.** Existen significativas diferencias en el espectro de las señales obtenidas de un sensor específico cuando el sistema opera a diferentes niveles de torque en la mismo tipo de defecto. En la Figura 45 se muestra esta situación para torques de 27% y 100% para una condición de no-defectos en el sistema. Con la comparación de esta graficas se demuestra que para cualquier tipo de defecto en el que se encuentre actualmente el sistema, existen pequeñas variaciones del espectro causadas por maniobras de vuelo o resistencia del viento que modifican el torque del mismo. Por lo

tanto en los datos obtenidos se cumple una propiedad, que al igual que en el reconocimiento de voz (para un fonema o silaba), es de naturaleza cuasi estacionaria.

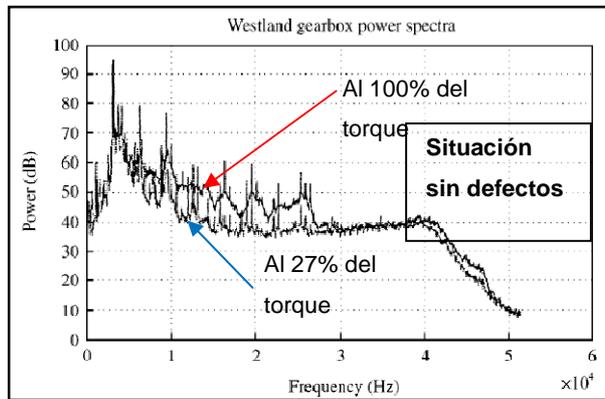


Figura 45. Condición del sistema sin defectos a 27% y 100% del torque nominal<sup>35</sup>.

**Situación 2.** Se manifiestan diferencias en el espectro de señales de un sensor específico cuando el sistema opera al mismo torque, pero en diferentes tipos de defectos. En la Figura 46 se muestra esta situación para un torque del 100% en la situación en donde no hay defectos y en la situación de desgaste del engranaje de espiral cónico de entrada (defecto # 4 de la tabla 1). Se demuestra con estas graficas que para un mismo torque de operación del sistema, las diferencias espectrales de dos situaciones defectuosas son bastante significativas teniendo un importante cambio en la banda de los 15 a 40 kHz.

Con esta situación se demuestra que la aplicabilidad del CBM en esta situación puede llegar a ser muy efectiva debido a que se facilita la diferenciación entre los tipos de defectos gracias a las relevantes discrepancias de sus espectros. Este hecho facilita la tarea de diagnostico de fallas al identificar una situación novedosa de falla de una situación no defectuosa.

<sup>35</sup> Imagen tomada de : BUNKS, Carey et. al. Condition-Based Maintenance of Machines using Hidden Markov Models. En: Mechanical Systems and Signal Processing. (march, 2000); p.601

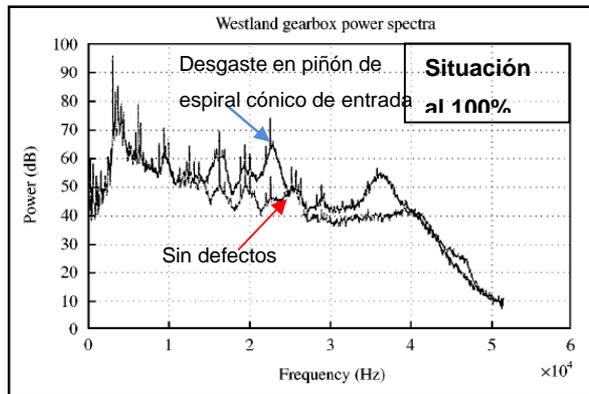


Figura 46. Condición del sistema al 100% del torque nominal, con y sin defectos<sup>36</sup>.

**Situación 3.** Se observan diferencias en el espectro de las señales de un sensor específico cuando el sistema se encuentra en un mismo tipo de defecto pero en diferentes niveles de severidad. En la Figura 47 se demuestra este hecho al comparar espectros de un tipo de defecto a un mismo torque, pero en los dos niveles de severidad (I y II). Se observa importantes diferencias en el espectro a diferentes niveles en una banda de frecuencia de los 15 a 40 kHz.

Con esta situación se demuestra que tal discrepancia de espectro para diferentes niveles de severidad permite el desarrollo del proceso de pronostico (estimación de la vida útil restante) en el sistema de transmisión del Westland.

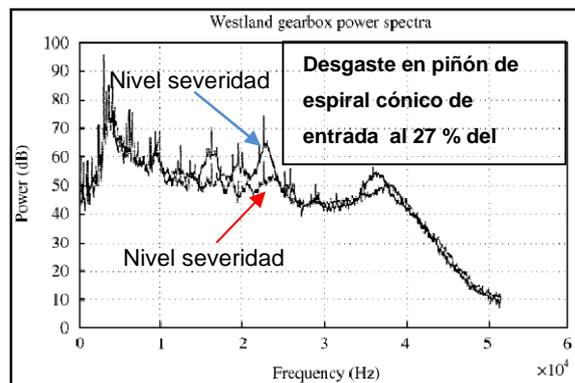
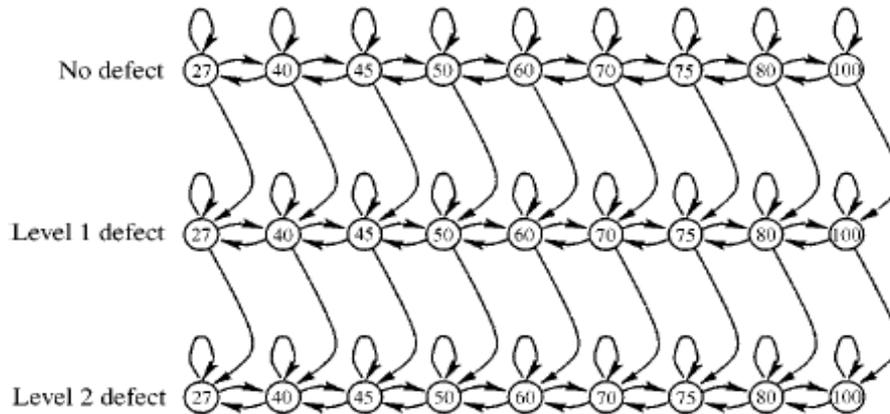


Figura 47. Condición del sistema a diferentes niveles de severidad a un mismo torque y defecto.

<sup>36</sup> Ibid.,p.602-603.

**4.3.2.2. Selección de estados y topología.** Al determinar la evaluación de 68 casos diferentes de condición (conformado por los pares torque-tipo de defecto, en diferentes niveles de severidad) del sistema de transmisión del helicóptero Westland, los autores determinan en construir un HMM cuyo número de estados corresponda al mismo número de condiciones a evaluar, que en este caso sería 68 estado. Cada uno de estos estados contiene características espectrales distintas entre sí, correspondientes a un porcentaje de torque, un tipo de defecto y un nivel de severidad del defecto. La topología utilizada para la interconexión de estos estados se observa en la Figura 48 en donde se representa la transición entre las condiciones del sistema de transmisión para un solo tipo de defecto.



**Figura 48. Topología de estados del sistema de transmisión del Westland para un solo defecto.**

En la primera fila se observa 9 estados relacionados cada uno a un porcentaje de torque cuando no existe defecto alguno. La segunda y tercera fila representa cada una los dos niveles de severidad para un solo tipo de defecto con sus respectivos 9 estados relacionados a un porcentaje de torque. Se interpreta de esta topología que en cualquier nivel de severidad (no- defectos, Level I o Level II), se pueden presentar transiciones entre torques con incrementos y descensos dependiendo de las condiciones de carga mecánica (fuertes vientos, maniobras) experimentado por las hélices y el resto de la transmisión. Se entiende además, que después de hacer una transición de un nivel de severidad inferior a otro superior, no existe transición de

regreso, indicando este hecho el proceso de deterioro gradual de los componentes monitoreados.

Una versión completa de la topología del HMM para esta aplicación se observa en la Figura 49 el cual contiene todos los 8 tipos de defectos. Esta topología es una extensión de la presentada en la Figura 48 en donde se repite la segunda y tercera fila para cada tipo de defecto y se conectan todas en paralelo con la primera fila (no-defectos). Con esta topología se puede hacer una transición desde una situación donde no hay defectos a cualquiera de los 8 tipos de defectos y sus niveles de severidad.

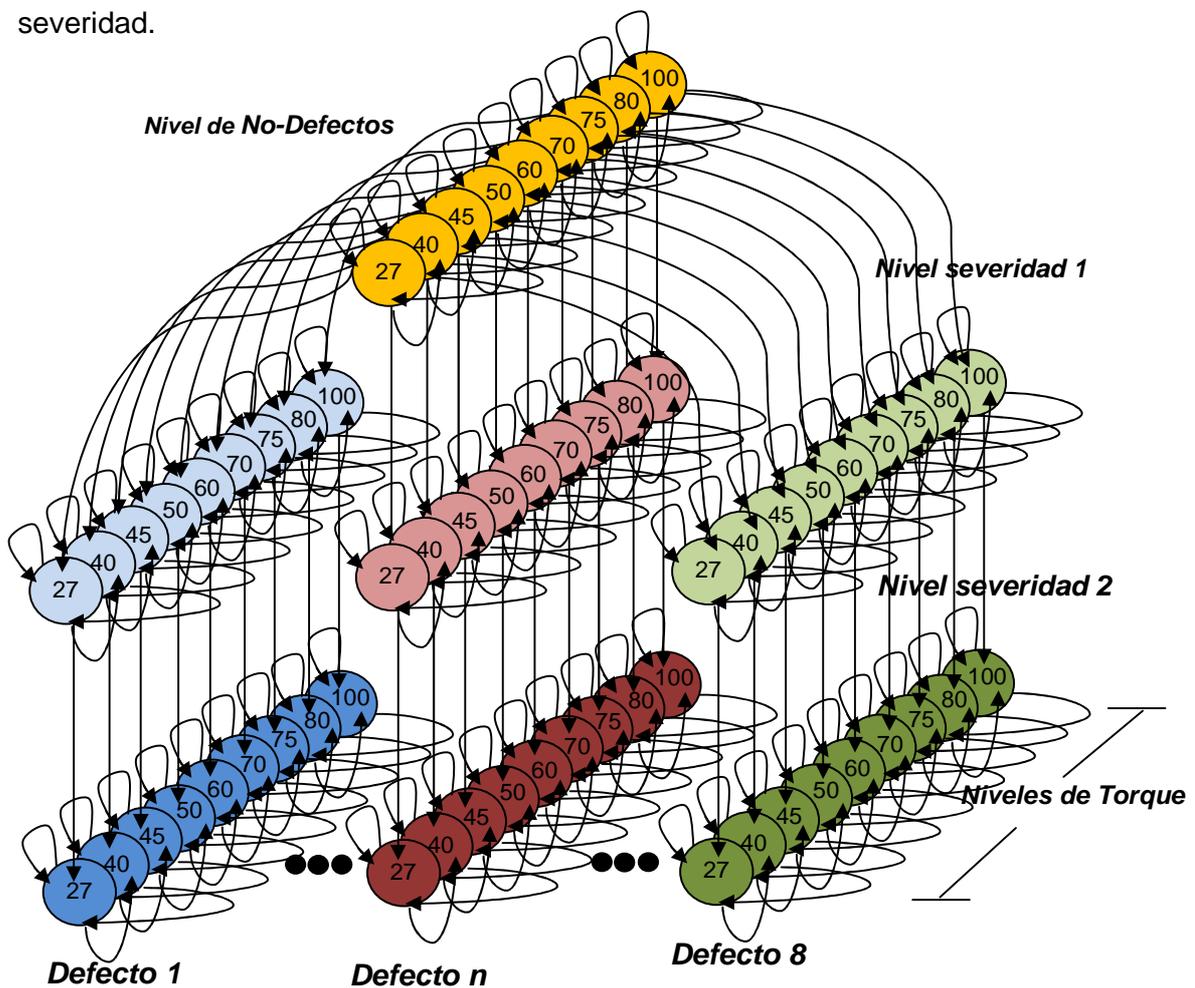
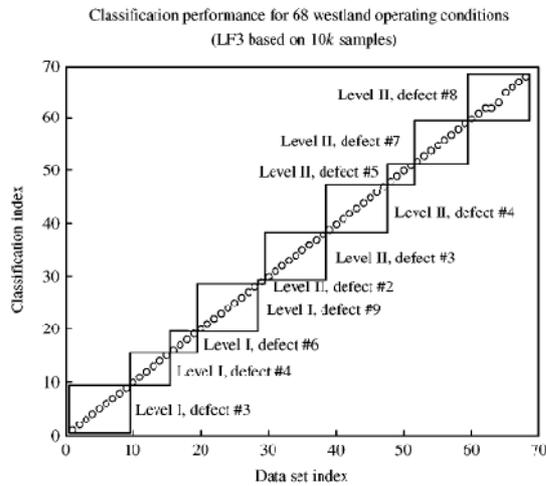


Figura 49. Topología de estados del sistema de transmisión del Westland para todos los defectos.

**4.3.2.3. Selección de Observaciones.** Los autores además se aseguran en diseñar un modelo de densidad probabilística para cada uno de los estados con el fin de hacer un posterior análisis de la habilidad de clasificación de distintas situaciones. A cada estado se le relaciona una función de densidad de probabilidad Gausiana de 8 dimensiones debido al número de sensores instalados en el sistema.

Como observaciones para cada función gaussiana se utiliza una transformación de los datos correspondientes a las vibraciones del sistema de transmisión de un defecto específico en coeficientes espectrales para cada sensor (dimensión), permitiendo la reducción de la cantidad de datos a ser procesados efectivamente en el proceso del HMM. La razón principal de esta transformación se debe al gran volumen de datos obtenidos de una tasa de muestreo de 100K muestras por segundo (frecuencia ideales de muestreo para ondas de vibración) para cada uno de los 8 sensores teniendo el HMM que enfrentar por ejemplo, a 800K muestras en un solo segundo (hasta millones de datos en unos cuantos segundos) si la transformación a coeficientes espectrales no se ejecutara. Es notorio por lo tanto que la transformación es necesaria para el procesamiento de muestras por parte del HMM.

**4.3.3. Entrenamiento y pruebas de clasificación de modelos.** Para medir la habilidad que tiene los 68 modelos Gaussianos (estados) de reconocer los datos pertenecientes a su propia clase de condición de operación en el Westland, estos fueron utilizados en experimentos de clasificación. Los experimentos se basaban en evaluar cual de los 68 modelos presentaba la máxima probabilidad sobre los datos observados de una condición específica del sistema de transmisión. En la Figura 50 se observa el resultado obtenido del experimento de clasificación para conjuntos de 10000 muestras por condición de operación, en donde el eje de las ordenadas muestra el índice del modelo gaussiano el cual clasificó con la mayor probabilidad, el conjunto de datos indexado para dicha condición que se muestra en el eje de las abscisas.



**Figura 50. Clasificación de condiciones de operación.**

En la figura se observan 68 círculos pertenecientes cada uno a un conjunto de datos clasificados en 68 modelos de condición de operación. Se observa además que estos círculos se encuentran agrupados en cuadros que corresponden a un mismo tipo de defecto en determinado nivel de severidad (I o II), como por ejemplo, el primer cuadro inferior izquierdo representa el defecto #3 en el primer nivel de severidad. En cada cuadro no hay más de 9 círculos, correspondiendo cada uno de ellos a un porcentaje del torque nominal. Se nota además que cada defecto y su determinado nivel de severidad, no es evaluado para todos los torques (defecto #6 en el primer nivel de severidad con solo 4 porcentajes de torque, etc.), e incluso, no todos los defectos se evalúan en los dos niveles de severidad (no se tiene en cuenta un segundo nivel de severidad para el defecto #6, ni un primer nivel de severidad para el defecto # 2, etc.).

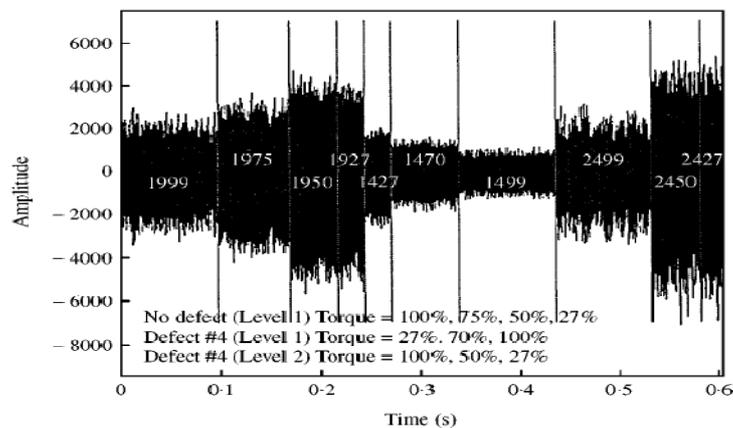
Una buena clasificación de todos los modelos se observa si todos los círculos están alineados en la diagonal de identidad. Se observa en este caso dos pequeños errores de clasificación en el nivel 2 del defecto # 8 donde el torque al 50% es mal clasificado al 45% y el 60% es mal clasificado al 50%.

Los autores realizan los mismos experimentos reduciendo el número de muestras a 1000 y 200, obteniendo un incremento de pequeños errores en clasificación pero que no

llegan a ser más que errores de clasificación entre torques dentro de un mismo defecto y nivel de severidad (un mismo cuadro).

Los autores realizan otro tipo de experimento para corroborar la habilidad de clasificación, esta vez, en un conjunto de datos provenientes de un sistema de transmisión del helicóptero Westland en funcionamiento. La diferencia de esta prueba con la anterior esta en probar los 68 modelos en tareas de clasificación e identificación de una *serie de diferentes condiciones* que se desarrollan como un proceso continuo en el tiempo, que es lo que normalmente ocurrirían en un sistema real, y no como una condición aislada y estática sin sucesión o precesión de otras condiciones de operación.

Las pruebas son ejecutadas usando un conjunto de datos compuestos en donde se juntan datos de 10 de las 68 diferentes condiciones de operación. Un ejemplo de este conjunto de datos se muestra para el sensor #7 (Eje dentado de babor) en la Figura 51, teniendo en cuenta que el experimento se hace para todos los 8 sensores simultáneamente (no solamente el sensor #7).

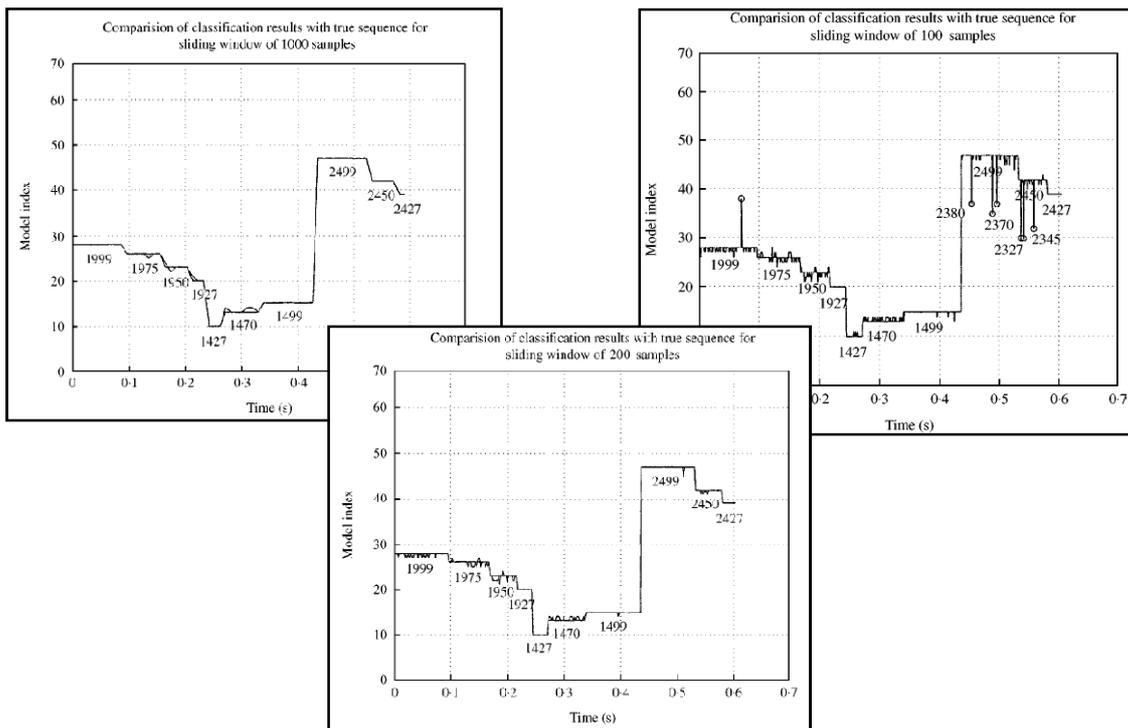


**Figura 51. Secuencia de datos para el sensor # 7.**

La figura muestra los 10 segmentos delineados por líneas verticales. Cada segmento contiene un código de 4 dígitos con el formato XYZZ, donde X se refiere a la severidad del defecto (1 o 2), Y se refiere al tipo de defecto (1-9), y ZZ representa el nivel del torque. El segmento que se encuentra más a la izquierda es un caso sin defectos a un

torque del 100% (1999). Seguidamente, hay una disminución del torque a niveles de 75 (1975), 50 (1950) y luego 27 (1927) en el cual hasta este punto se introduce el defecto # 4 (desgaste del piñón de espiral cónico de entrada) con el mismo torque de 27% (1427). Luego se incrementa el nivel de torque a 70% (1470) y luego a 100% (1499) con el mismo defecto (#4). Posteriormente, la situación se vuelve más seria cuando el defecto #4 evoluciona al segundo nivel de severidad al 100 % del torque (2499). El nivel de torque finalmente decrece a 50 % (2450) y luego a 27% (2427) para este mismo defecto.

Para identificar la secuencia de datos mostrada en la figura anterior, se hacen pruebas con ventana de datos corridizas sin traslape utilizando 1000, 200 y 100 muestras sobre los datos. Los resultados de estas pruebas se observan en la Figura 52.



**Figura 52. Identificación secuencia de condiciones con diferentes números de muestras<sup>37</sup>.**

<sup>37</sup> Ibid., p.607-608

En esta figura se comparan a través del tiempo las condiciones reales de operación caracterizadas por los datos de prueba con las condiciones de operación clasificadas por los modelos. La codificación de las condiciones es la misma manifestada para la Figura 51. Se observa en la grafica izquierda (con ventanas de 1000 muestras) pequeñas imprecisiones de clasificación desde 0.1 a 0.35 segundos, que no son más que errores de clasificación de torques. En la grafica central (con ventanas de 200 muestras) se detalla imprecisiones más relevantes que la grafica anterior pero que permanecen siendo solo discrepancias entre torques. Finalmente en la grafica de la derecha (con ventanas de 100 muestras) se observa imprecisiones importantes de mala clasificación entre defectos y torques.

En las dos pruebas anteriores se comprueban que los modelos utilizados son robustos para la clasificación de defectos e incluso de torques contribuyendo a el desarrollo de la etapa del diagnostico de fallas. Además se comprueba que el uso de un número alto de muestras mejora la eficacia de la clasificación.

**4.3.4. Diagnostico y Pronostico con modelos HMM.** En la sección anterior se mostraron las pruebas de clasificación en los 68 modelos en forma independiente y en forma secuencial, colocando arbitrariamente una serie de condiciones de operación con tiempos irreales de duración estipulados de forma deliberada.

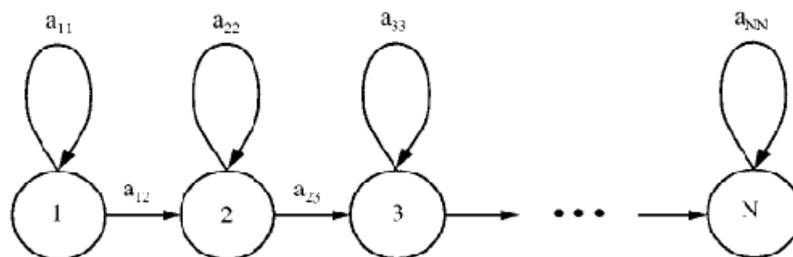
El diagnostico de condiciones de operación de un sistema real requiere que estos modelos tengan restricciones de transición entre ellos, tiempo de duración y frecuencia de ocurrencia facilitando las tareas de detección y estimación de condiciones de una manera diferente a la propuesta en la sección anterior.

La selección de los valores de densidad de la transición entre estados (modelo o condiciones) se hace en base a información adicional que se tenga acerca de conocer las frecuencias relativas de visita de cada estado a partir de otro, y de modelos explícitos que contienen el conocimiento de las características físicas de varios componentes de la maquina. Mas explícitamente, la selección de la matriz de transiciones se basa en información previa acerca de la frecuencia de ocurrencia de cada defecto y el tiempo promedio utilizado para operar en cada nivel de torque. Al

agregar una matriz de transiciones a los modelos definidos en la sección anterior, se está describiendo el comportamiento real del sistema de transmisión de Westland como una secuencia de estos modelos en tiempos reales probables.

Por otro lado, como función esencial del CBM está la planeación de mantenimiento con ayuda de resultados arrojados por una etapa de Pronóstico que calcula el tiempo de vida útil restante de componentes de alta criticidad de una maquina. Como se comprobó en la situación 3, la tarea de pronóstico puede ser llevada a cabo gracias a las diferencias que se presenta entre los espectros de los datos obtenidos en los sensores cuando operan en diferentes niveles de severidad para un mismo tipo de defecto.

En la Figura 53 se muestra un HMM que puede ser utilizado para estimar el tiempo de vida útil restante promedio del sistema de trasmisión. En esta figura, cada estado representa el grado de severidad de un tipo de defecto particular del sistema de trasmisión. El proceso aleatorio de cada estado es determinado por una fila de la Figura 48. Es decir, que el estado 1 de la Figura 48 está asociado a un proceso aleatorio generado por la fila de NO defectos de la misma figura, el segundo estado es asociado al proceso aleatorio de la fila de primer nivel de severidad de la Figura 48, el tercer estado al nivel segundo nivel de severidad, y así sucesivamente (si existiese otros niveles de severidad).



**Figura 53. Topología de estados para el pronóstico del sistema<sup>38</sup>.**

---

<sup>38</sup> Ibid., p.610.

Para cada estado  $n$  de este HMM, hay una probabilidad de permanecer en ese mismo estado de  $a_{nn}$ , y una probabilidad de dejarlo de  $a_{nn,n+1}$ . Debido a que los defectos siempre avanzan en deterioro y nunca en mejora, se utiliza una topología de izquierda – derecha que permite transiciones solo en la dirección de deterioro.

Para el pronóstico de condiciones, se asume que el último estado  $N$  de la cadena representa el estado en donde no hay tiempo de vida útil restante, pues este estado representa un nivel de alto deterioro e inutilidad del equipo. Por lo tanto, el tiempo promedio  $t^*$  para llegar a este estado  $N$  a partir de cualquier estado de nivel de severidad  $n$ , se calcula como una función del número medio de paso requeridos para ir del estado  $n$  a  $N$ . Esto se lleva a cabo calculando la probabilidad  $p_k$  de ir del estado  $n$  al estado  $N$  en exactamente  $k$  pasos y luego hallando la suma

$$t^* = \sum_{k=1}^{\infty} k p_k$$

Para que el problema de pronóstico se lleve en forma optima, se debe hacer una selección apropiada de las probabilidades de transición. Estos se determinan a partir de conocimiento experimental proporcionada por los fabricantes de componentes, por ejemplo, del número de operaciones límite antes de fallar de un componente a determinadas condiciones, el número de horas máximo de operación, el numero de millas de operación, etc. Todo este tipo de información se encuentra por lo regular para varios tipos de componentes y puede ser incorporado dentro del HMM de la Figura 53 como probabilidades de transición.

### RESUMEN TÉCNICAS PREDICTIVAS EN MANTENIMIENTO DE EQUIPOS: DIAGNOSTICO Y PRONOSTICO CON HMM'S

#### 1. Naturaleza y dificultades del reconocimiento condiciones en equipos.

Comprende la ejecución de tareas de diagnostico y pronostico para la determinación de condiciones de operación de un equipo de terminado bajo el esquema de Mantenimiento Basado en Condición (CBM). Se fundamenta bajo la suposición de que ningún equipo falla repentinamente, sino que por el contrario, este presenta una evolución temporal de deterioro en donde los estados iniciales del proceso no son perceptibles por el ser humano. Un HMM implementado sobre un monitoreo de condición ayuda a establecer el estado de deterioro a través de medidas de sensores (vibración, calor, presión, etc).

#### 2. Estructura de HMM.

- **Lo que representa un HMM:** Representa una maquina de transición entre condiciones de deterioro de un equipo, iniciando por un estado de funcionamiento optimo hasta un estado de deterioro total. Describe temporalmente el comportamiento de deterioro de un equipo determinado.
- **Topología:** Dependiendo de la cantidad de operaciones a detectar y la relación entre tales estados se puede establecer una topología cualquiera. Sin embargo, el proceso de deterioro de un equipo presenta estados irreversible por lo que una topología de izquierda –a derecha es adecuada para su modelamiento.
- **Numero de estados:** Depende de la cantidad de condiciones de operación que se desee evaluar en el equipo.
- **Observaciones O:** La selección de las observaciones depende de las variables medidas por sensores para determinar la condición de un equipo. Por lo general a estas señales se les aplica un pre-procesamiento para determinar las características en frecuencia o tiempo de tal señal. La distribución de estas observaciones sobre los estados depende de las características representativas de estas señales a una condición de operación determinado.

#### 3. Entrenamiento del modelo de deterioro con HMM's.

Al sistema debe se le debe proporcionar información de varias condiciones de operación para que cada estado del sistema asocie su estado de deterioro. Se hacen pruebas de clasificación de condiciones para corroborar que cada condición sea detectada por un estado particular.

#### 4. Diagnostico y pronóstico con HMM's.

Entrenado el modelo HMM, se puede determinar es estado de deterioro actual de una equipo evaluando las señales emitidas. Con este ejercicio se diagnostica el estado de operación actual del sistema.

El HMM entrenado permite además pronosticar la probabilidad de falla y el tiempo de vida útil restante del equipo antes de someterse a un estado de deterioro total. Gracias a la conexión de estado por niveles de severidad, se puede determinar un estimativo de tiempo restante.

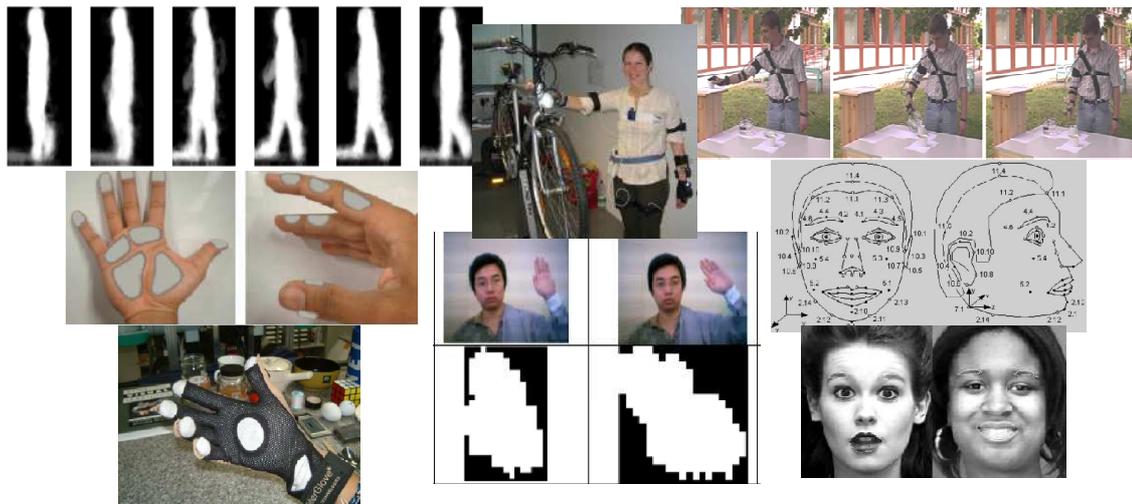
### 4.4. APRENDIZAJE Y RECONOCIMIENTO DE PATRONES DE MOVIMIENTO CORPORAL EN SERES HUMANOS

El cuerpo humano está dotado para realizar múltiples actividades de movimiento como respuesta a una estrategia mental previamente planeada para su ejecución. Algunas de estas actividades corresponden a necesidades de comunicación no verbal para la transmisión de un mensaje, sea con gestos faciales o ademanes que manifiestan emociones o estados de comportamiento. Otras actividades corresponden a la necesidad de ejecutar un trabajo específico, como por ejemplo levantar y desplazar un objeto, caminar, correr, ensamblar piezas, etc., donde no necesariamente se requiera transmitir un mensaje en forma intencional.

Sin embargo, cualquiera que sea el tipo de actividad ejecutada por el ser humano, el factor común predominante radica en que cada una de ellas presenta un significado único que se distingue de las demás y que intencionalmente o no, transmiten por parte del que la ejecuta un mensaje para aquellos que la perciben. Esto abre la posibilidad de establecer una interacción comunicativa entre dos personas únicamente con las expresiones de su cuerpo en donde el esquema acción- respuesta es su fundamento.

**4.4.1. Avances en reconocimiento de patrones de movimiento del cuerpo humano.** Uno de los intereses de la inteligencia artificial consiste en poder crear modelos que reconozcan comportamientos humanos interpretando los movimientos que estos hacen con algunas partes de su cuerpo. Algunas ramas que han seguido esta idea (Figura 54) se reflejan en trabajos presentados como en [16], en donde se hace el reconocimiento de emociones por medio de la extracción de características con el Análisis de Componentes Principales (PCA) y Cadenas Ocultas de Markov en expresiones faciales. Similarmente, en [3] y [10] se desarrolla el reconocimiento de una secuencia de gestos manuales (Hand Gesture) con datos obtenidos con guantes de sensores táctiles o imágenes utilizando Cadenas Ocultas de Markov para su identificación. En [23] se hace mención a una aplicación particularmente interesante en el que consiste en el reconocimiento de personas a través de su forma de caminar con

un esquema que asume que durante la caminata la persona hace una transición entre una secuencia de varias posturas discretas. En [11] se hace el reconocimiento de movimientos manuales dependientes del usuario cuando escribe (Handwriting recognition) haciendo uso de Cadenas Ocultas de Markov discretas y análisis de vectores de aceleración tridimensional. En [4] se hace la clasificación y reconocimiento de fases de movimiento de seres humanos en actividades como levantar una jarra con liquido, servir en un vaso el contenido de la jarra, revolver el contenido con una cuchara, etc., utilizando técnicas como las Redes Neuronales y las HMM's. También en [22] se presenta el reconocimiento de actividades continuas basada en datos obtenidos de la posición de las manos y brazos de la persona que lleva a cabo actividades típicas de reparación de una bicicleta y comprobar que la secuencia de acciones sea adecuada para la ejecución del proceso de reparación.

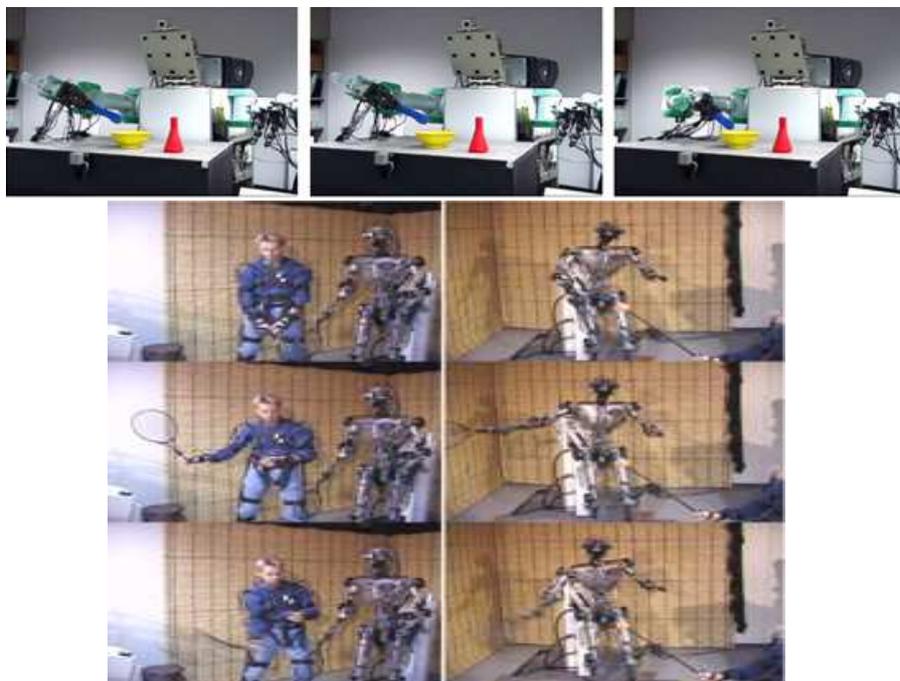


**Figura 54. Reconocimiento de patrones de movimiento corporal<sup>39</sup>.**

Otras aplicaciones utilizan el mismo principio de interpretación o reconocimiento de movimiento en seres humanos para someter a una maquina a realizar las mismas actividades bajo el esquema de aprendizaje por imitación (Figura 55). En [14] se utiliza

<sup>39</sup> Imágenes tomadas de: [1616],[3],[10], [23], [4] y [22].

un robot para la enseñarle tareas de manipulación del día a día por parte de cualquier tipo de usuario por medio de imágenes de video bajo el esquema de Aprendizaje a partir de Observaciones (Learning from Observation). De manera similar en [21] se examina desde el punto de vista puramente computacional el problema de imitación, enfatizando en el lado generativo de la imitación (control motriz), asumiendo que un sistema perceptual ya ha hecho la identificación de características importantes de un movimiento demostrado.



**Figura 55. Ejecución de tareas por método de imitación<sup>40</sup>.**

Algunas aplicación más complejas requiere que un sistema interactúe en línea con un ser humano, requiriendo además de interpretar y reconocer lo que el humano hace, responder con movimientos acordes al estímulo dado. Esto requiere de modelos más sofisticados donde se tenga en cuenta las interacciones que tendrían convencionalmente dos seres humanos, requiriendo de algún modo que utilizar un método de demostración de tales interacciones. En [24] se observa el caso donde se

---

<sup>40</sup> Imágenes obtenidas de: [1414], y [21].

aplica el proceso de interacción entre un humano y un robot, donde este último aprende el proceso a partir de una demostración de interacción humano-humano.

En esencia, los modelos utilizados en las aplicaciones mencionadas anteriores deben proporcionar una descripción de los movimientos característicos del cuerpo humano como un proceso que se desarrolla en el tiempo pues son estas una secuencia de maniobras que se llevan a cabo hasta dar por terminado la ejecución de la tarea.

Este esquema justifica la utilización de HMM's como herramientas útiles para el modelamiento de patrones de movimiento en seres humanos. La identificación del doble proceso estocástico de esta tarea se refleja en un proceso observable que equivale a la secuencia de datos de las posiciones del cuerpo y a un proceso oculto que representa la intención, mensaje o significado de la serie de posiciones ejecutadas.

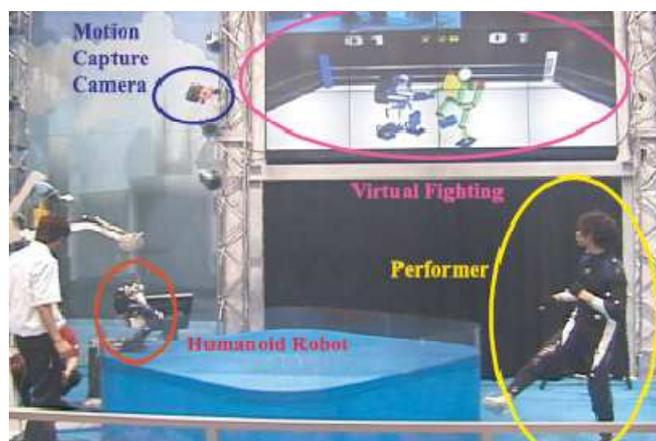
A continuación se identifica la estructura de una Cadena Oculta de Markov para el reconocimiento de patrones de movimiento en un proceso de interacción entre dos seres humanos.

**4.4.2. Selección de la estructura  $\lambda$  de HMM para reconocimiento de patrones de movimiento corporal.** Para hacer una descripción de un caso real donde se haga reconocimiento de movimientos de seres humanos, se tomará como ejemplo el proyecto desarrollado por Wataru Takano et al. [24] cuyo documento se titula "*Primitive Communication based on motion recognition and Generation with Hierarchical Mimesis Model*"<sup>41</sup> en el cual modelan procesos de comunicación no verbal entre dos personas utilizando tres grupos de Cadenas Ocultas de Markov jerárquicamente combinadas para reconocer el movimiento de un humano y generar el movimiento de interacción o respuesta por parte de un robot.

---

<sup>41</sup> Todas las imágenes que se presentan en esta aplicación fueron tomadas de este artículo.

El escenario de prueba utilizado por los autores consiste en la interacción de dos personas en una pelea de Kick Boxing en el cual se analizan los patrones de interacción de defensa de uno de ellos como respuesta al ataque del otro contrincante (Figura 56). Estos patrones de interacción entre los dos combatientes son aprendidos por un robot con la demostración dada por los humanos, que posteriormente son utilizados en un escenario de pelea humano - robot donde este último generará movimientos de defensa acordes a los ataques dados por su humano contrincante.

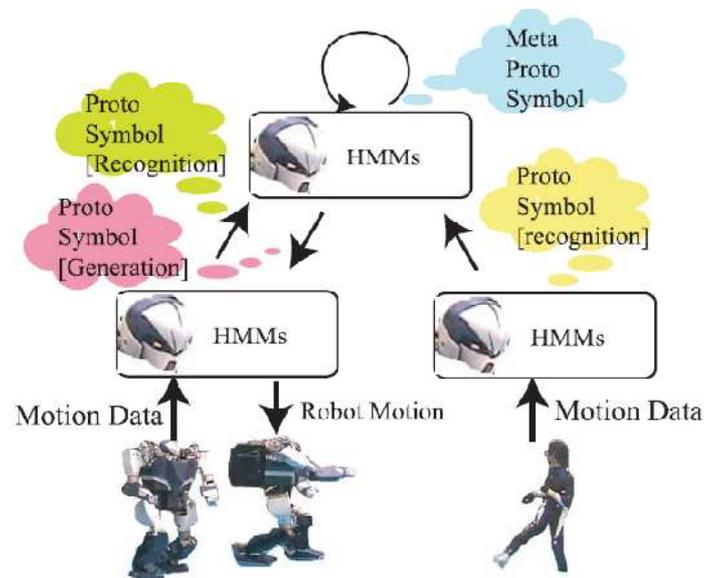


**Figura 56. Escenario de KickBoxing entre humano y robot.**

Para llevar a cabo este tipo de interacción, el robot debe como primera medida reconocer los patrones de movimiento de su contrincante para luego asociarlo con el aprendizaje obtenido en la demostración a un movimiento de interacción defensiva adecuado que generará los comandos de control a la estructura física del robot.

Los autores diseñan un *modelo de comunicación mimética* que se compone de dos tipos de HMM's jerárquicamente estructurados, en donde los HMM's pertenecientes a una capa inferior modela los patrones de movimientos básicos de cada contrincante, y un HMM de capa superior que modela los patrones de interacción entre los dos protagonistas para representar en forma adecuada sus comportamientos regulares de acción - reacción. Los autores denominan a los HMM's de capa inferior como *proto-*

símbolos y a los de capa superior como *meta-proto-símbolos*. En la Figura 57 se observa el modelo de comunicación mimética diseñado por los autores.



**Figura 57. Modelo de comunicación mimética.**

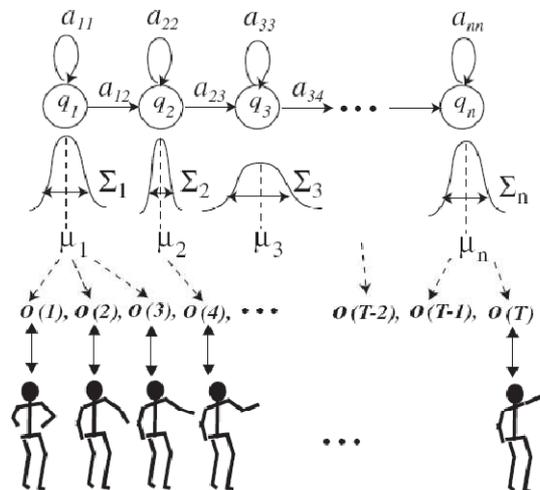
Se observa que los HMM's de la capa inferior obtienen los datos de movimiento del robot y el humano para reconocer aquellos proto-símbolos que con mejor probabilidad represente los datos obtenidos para cada uno de ellos. Luego que los proto-símbolos de ambos protagonistas sean identificados, estos entran a los HMM's de capa superior para encontrar aquel HMM que represente el meta-proto-símbolo que con mayor probabilidad modela la interacción entre el robot y el humano a partir de sus proto-símbolos reconocidos (más adelante se describirá la forma en que los autores correlacionan los proto-símbolos para una meta-proto-símbolo dado). Por último, se selecciona el proto-símbolo que con más alta probabilidad represente en forma adecuada el movimiento que el robot debería ejecutar para responder a los estímulos dados por el humano.

**4.4.2.1. Descripción del dominio de proto-símbolos y domino de meta-proto-símbolos.** El conjunto de proto-símbolos y meta-proto-símbolos seleccionados por los autores para probar su modelo de comunicación mímica son los siguientes:

- Proto-símbolos: Se seleccionaron 8 tipos de patrones de movimientos entre ellos: Agacharse, patada izquierda, protección con la piernas izquierda, puñetazo izquierdo, patada derecha, protección de la pierna derecha, puñetazo derecho y posición de pelea.
- Meta-proto-símbolos: Teniendo en cuenta que estos están compuestos por proto-símbolos de cada uno de los contrincantes, algunos de los meta-proto-símbolos son:
  - Uno de los combatientes da una patada izquierda (derecha) y luego se protege con la pierna derecha (izquierda) como respuesta a la patada izquierda (derecha) del otro.
  - Uno de los combatientes proporciona un puñetazo izquierdo (derecho) y luego se agacha como respuesta al puñetazo izquierdo (derecho) del otro.

**4.4.2.2. Topología, estados y observaciones de proto-simbolos y Meta-proto-simbolos.** En la Figura 58, se observa la representación de los patrones de movimiento que se desarrollan en el tiempo por medio de un HMM con topología izquierda- derecha, junto con sus funciones de probabilidad y sus respectivas observaciones  $O(t)$ .

**Figura 58. Representación de patrones de movimiento (protosimbolos) con HMM.**



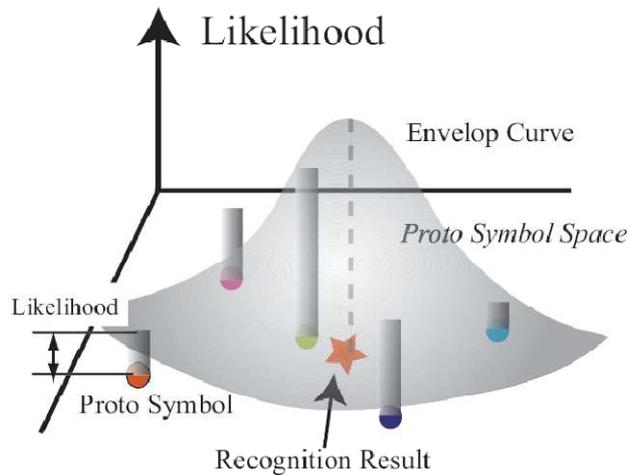
Se ilustra que un proto-símbolo particular (por ejemplo, dar un puñetazo con la mano izquierda) es interpretado como una secuencia de movimientos que se desarrollan en función del tiempo. Cada observación  $\mathbf{O}(t)$  o captura instantánea de la posición del cuerpo, está definido probabilísticamente dentro de un estado específico del HMM por medio de una función de densidad de probabilidad (pdf por sus siglas en inglés) de tipo gaussiano descrito por su valor medio  $\mu_n$  y su covarianza  $\Sigma_n$ . Las dimensiones de esta función dependen del número de variables necesarias para representar un proto-símbolo.

Las observaciones utilizadas para representar los patrones de movimiento son obtenidos por la medición de las coordenadas cartesianas de 25 marcadores que se encuentran adheridos al cuerpo del combatiente usando un sistema comercial de captura óptica de movimiento. Estas coordenadas son convertidas en medidas de ángulos de articulación, altura, inclinación, y nivel de balance del cuerpo humano, junto con sus velocidades. El tiempo de muestreo para la captura de movimientos es de 30 ms.

Por otro lado, la selección de la estructura HMM correspondientes a los meta-proto-símbolos es más compleja porque debe tener en cuenta que las interacciones (meta-proto-símbolos) que se desarrollan en el tiempo están conformadas por pares de proto-símbolos que cambian también con el tiempo. Para esto, los autores modelan las interacciones entre los dos combatientes creando un espacio de proto-símbolos por cada combatiente para representar las relaciones que existen entre ellos mismos.

Después de construir los espacios, se obtiene los patrones de interacción entre ellos convirtiendo los datos de movimiento en una trayectoria dentro del espacio. La obtención de los patrones de interacción se consigue calculando la probabilidad de que cada proto-símbolo sea capaz de generar los datos de movimiento observados contenidos en una ventana de datos de ancho  $w_{span}$ . Luego computa la localización del movimiento observado en el espacio de proto-símbolos a partir de la probabilidad calculada anteriormente. La localización del movimiento observado se describe por el

centro una función Gausiana que describe las probabilidades de este movimiento con cada proto-símbolo. En la Figura 59 se muestra la localización de un patrón de movimiento descrito por las observaciones  $\mathbf{O}(t)$ , dentro de un espacio de proto-símbolos.



**Figura 59. Espacio de protosímbolos.**

En este espacio, se observan varios proto-símbolos entrenados previamente (esferas de colores), a los cuales se les calcula un valor de probabilidad con respecto a una curva gaussiana envolvente, cuyo centro (estrella naranja) es el resultado de reconocimiento de un patrón de movimiento descrito por las observaciones  $\mathbf{O}(t)$ .

Cuando un proceso de interacción se lleva a cabo, una secuencia de patrones de movimiento deben ser ubicados dentro del espacio de proto-símbolos obteniendo una trayectoria continua de símbolos (varias estrellas naranja) caracterizados por los centros o medias de las curvas gaussianas envolventes de cada movimiento descrito. Cada una de estas medias dentro del espacio de proto-símbolos representa la secuencia de observaciones  $\mathbf{x}_p(t)$  del contrincante  $\mathbf{p}$  que es utilizado por los HMM's correspondientes a los meta-proto-símbolos.

En la Figura 60 se muestra un HMM perteneciente a un meta-proto-símbolo que modela una interacción entre dos humanos en combate.

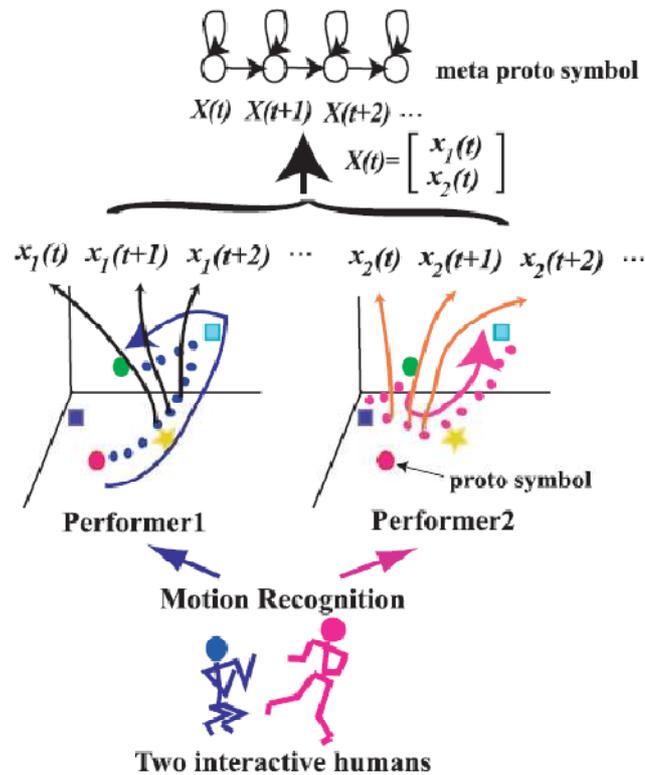


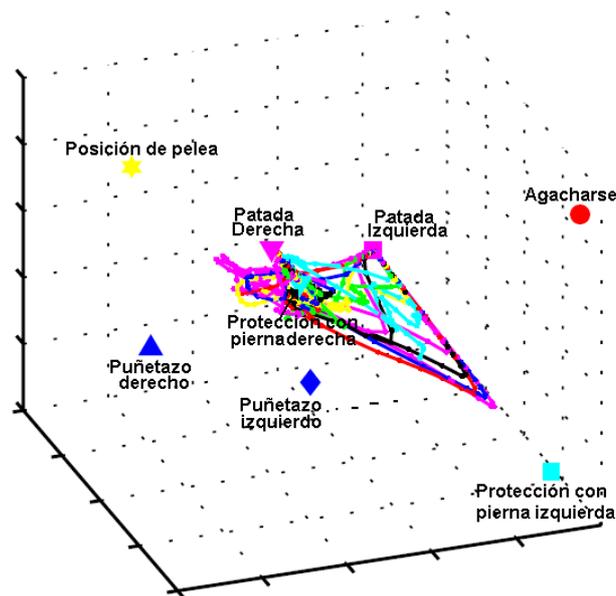
Figura 60. Representación de patrones de Interacción (meta-protosímbolos)

Se ilustra que cuando dos humanos interactúan, los patrones de movimiento de cada contrincante son ubicados dentro de su correspondiente espacio de proto-símbolos obteniendo una serie de puntos  $\mathbf{x}_p(\mathbf{t})$  ( $x_p(t)$ ,  $x_p(t+1)$ ,  $x_p(t+2)$ ,...) pertenecientes a trayectoria de símbolos desarrollados por cada humano  $\mathbf{p}$  (flecha curvadas de color azul y roja). La serie de puntos  $\mathbf{X}(\mathbf{t})$  representa las observaciones (en analogía a las observaciones  $\mathbf{O}(\mathbf{t})$  de los proto-símbolos) del comportamiento simultáneo de los dos combatientes a través del tiempo para los HMM's que modelan meta-protosímbolos.

**4.4.3. Entrenamiento de los modelos, Reconocimiento y Generación de movimiento.** Como ya se había mencionado, el aprendizaje de patrones de movimiento e interacción se realiza a partir de la demostración de dos combatientes humanos que ejecutan movimientos de ataque y defensa característicos de una pelea de kickboxing. El combate de un humano y un robot por lo tanto, se lleva a cabo después de que el sistema aprenda los patrones de movimiento e interacción de la demostración anterior.

La etapa de aprendizaje de un proto-símbolo consiste en optimizar los parámetros  $\lambda = \{A, B, \pi\}$  del HMM por el algoritmo de Baum Welch de manera que la probabilidad de que ese HMM genere los patrones de movimiento observados llegue a ser máxima. Se requiere por lo tanto varias secuencias del mismo movimiento (por ejemplo, agacharse) para obtener valores apropiados de los parámetros. El tiempo de muestreo es de 30 ms para la captura del movimiento y el número seleccionado de vectores dimensionales es de 46 (1.38 s). Luego de que los proto-símbolos estén definidos, se crea el espacio de proto-símbolos para cada combatiente como se mencionó anteriormente.

Los autores realizan una prueba para investigar la validez del reconocimiento de movimiento de varias interacciones utilizando el espacio de proto-símbolos de cada combatiente en forma independiente. Para esto adoptaron las secuencias de patrones de movimiento interactivo que se mencionaron la sección 4.4.2.1. El ancho de la ventana de datos de movimiento para el reconocimiento es de 180 ms, lo que significa que los datos de movimiento incluyen solo 6 tramas de datos capturados. En la Figura 61 se observa los resultados obtenidos por los autores en el espacio de proto-símbolos de uno de los contrincantes graficando 10 trayectorias de símbolos correspondientes a 10 repeticiones de una misma interacción cuando un combatiente propina una patada con la pierna izquierda y posteriormente se protege con la pierna derecha de la patada que le devuelve su oponente con la pierna izquierda.



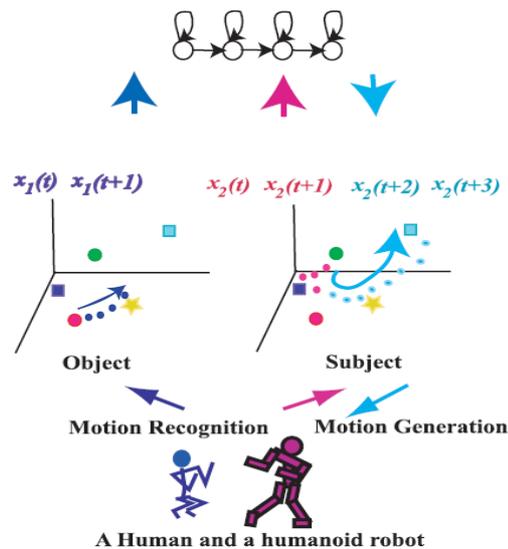
**Figura 61. Trayectoria de símbolos resultantes de una interacción.**

Se muestra que las trayectorias de símbolos tienden a comenzar en el proto-símbolo “patada izquierda” y posteriormente busca el proto-símbolo “protección con la pierna derecha” a través del proto-símbolo “protección con la pierna izquierda”. La causa de esta transitoria desviación se debe a que los patrones de movimiento de protegerse con la pierna izquierda y de dar una patada con esta misma pierna son similares. Estas pruebas se hacen para todas las interacciones y graficando las trayectorias en los espacios de proto-símbolos de cada combatiente.

Posteriormente estas trayectorias de símbolos en el espacio de proto-símbolos en la interacción de dos humanos son adoptadas como señales supervisadas para el entrenamiento de los HMM's de capa superior que corresponde a los meta-proto-símbolos a modelar.

**4.4.4. Generación de movimiento a través de meta-proto-símbolos.** Luego de que los parámetros  $\lambda$  de los HMM's de capa inferior y de capa superior hayan sido optimizados con el entrenamiento a través de la demostración de dos humanos, los autores demuestran la efectividad de su modelo mimético de comunicación para un robot que interactúa con un ser humano en un ámbito de combate. Para esto los patrones de movimiento de uno de los combatientes son medidos con el sistema de captura de movimiento. Cada secuencia de movimiento y cada patrón de interacción están relacionados en tiempo real con un proto-símbolo y un meta-proto-símbolo respectivamente.

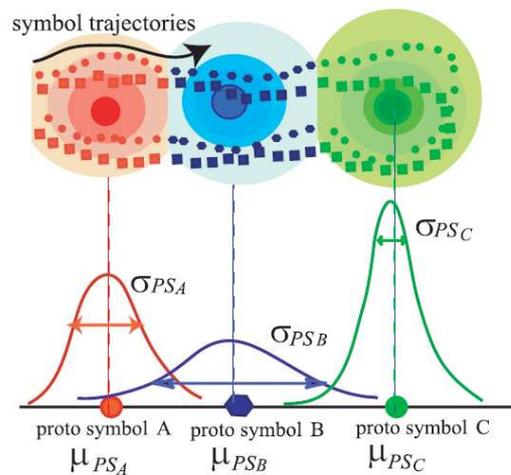
La fase de generación de movimiento para un robot humanoide se observa en la Figura 62.



**Figura 62. Fase de Generación de movimiento para un robot.**

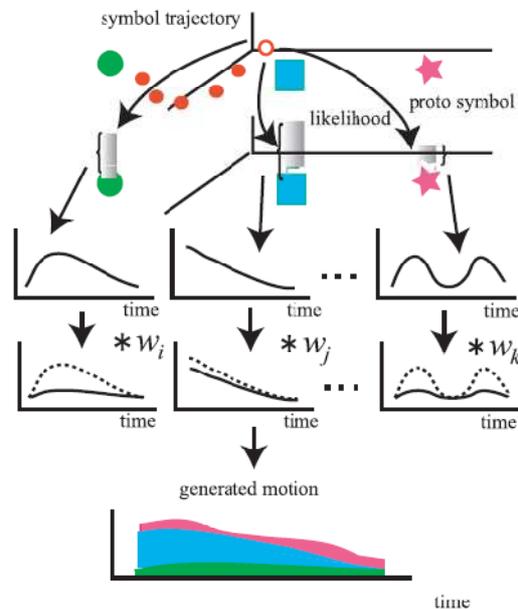
Como primera medida, el sistema reconoce el movimiento del humano y del robot y los ubica en su respectivo espacio de proto-símbolos. Luego a partir de las trayectorias parciales de cada uno (para el humano, los puntos azules  $x_1(t)$ ,  $x_1(t+1)$  y para el robot, los puntos rojos  $x_2(t)$ ,  $x_2(t+1)$ ), el sistema encuentra el meta- proto símbolo que tenga la mayor probabilidad para generar dichas trayectorias simultáneamente. Después de

seleccionado el meta-proto-símbolo más probable, se convierte a este en un generador de trayectoria de símbolos (para el robot flecha azul claro con los puntos  $x_2(t+2)$ ,  $x_2(t+3), \dots$ ) que plantearan un hipotético patrón movimiento que tendrá que hacer el robot como respuesta a los movimientos de su oponente. Luego de localizar estos puntos hipotéticos en tiempos futuros, se calcula la probabilidad de ellos con respecto a cada uno de los proto-símbolos debido a que son estos los que poseen la información de los patrones de movimiento que el robot debe ejecutar. El cálculo de esta probabilidad se hace con base en la creación de una función gaussiana para cada uno de los proto-símbolos teniendo estos su valor medio  $\mu_{PS_j}$  y su desviación estándar  $\sigma_{PS_j}$ . En la Figura 63 se observa como la trayectoria de símbolos hipotéticos se desplaza por las zonas predominantes de cada proto-símbolos dentro del espacio.



**Figura 63. Trayectorias de símbolos hipotéticos.**

Es posible por lo tanto evaluar para cada punto de esta trayectoria de símbolos la probabilidad de que este pueda ser generado por los proto-símbolos del espacio. Finalmente la generación de movimiento está caracterizado por la secuencia de observaciones  $\mathbf{O}_G(\mathbf{t})$  obtenidas a partir de una superposición de patrones de movimiento  $\mathbf{O}_{G_j}(\mathbf{t})$  que son proporcionadas por cada proto-símbolo y valoradas cada una proporcionalmente por un factor de peso  $W_j(\mathbf{t})$ . Este valor es proporcional a la probabilidad calculada de cada punto para cada proto-símbolo como se mencionó anteriormente. En la Figura 64 se ilustra este hecho.

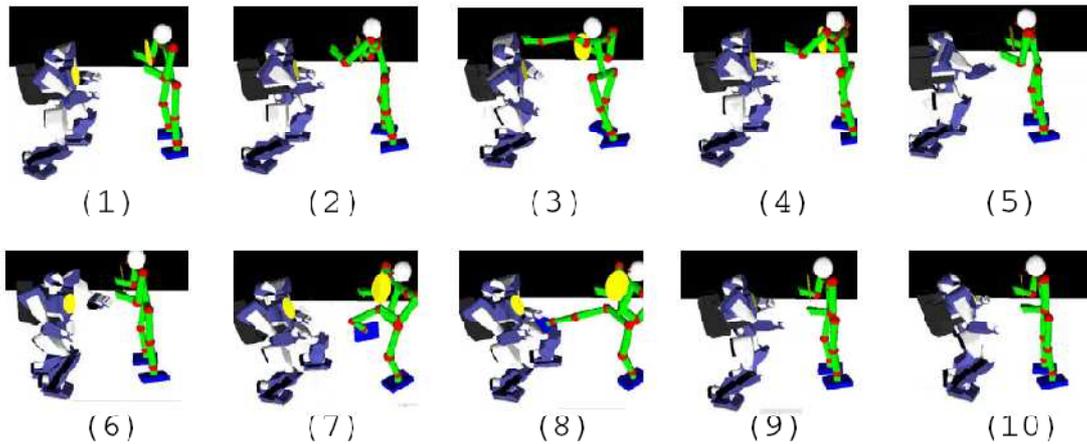


**Figura 64. Generación de movimiento por superposición de protosímbolos.**

Se detalla que uno de los puntos pertenecientes a una trayectoria se le evalúa un valor de probabilidad para cada proto-símbolo (círculo verde, cuadro azul y estrella roja). Cada proto-símbolo genera una secuencia de observaciones en el tiempo que es multiplicado por su factor de peso correspondiente producto de la probabilidad calculada. Como resultado se obtiene una secuencia de observaciones a partir de la superposición de las observaciones ponderadas de cada proto-símbolo (se muestra que las observaciones más predominantes son aquellas generadas por el proto-símbolo de cuadro azul).

En forma práctica, los autores primero ponen a prueba su modelo de comunicación mimética con la generación de patrones de movimiento por parte de un robot humanoide virtual de manera que este pueda reaccionar en respuesta a los movimientos del ser humano. Esto lo hacen con el fin de confirmar el funcionamiento del modelo en el ámbito computacional antes de implementarlo físicamente sobre los comandos que controlan los movimientos del robot. En la Figura 65 se muestra una

pelea virtual entre un robot humanoide y un humano, en donde la secuencia 1-5 se puede ver como el robot se agacha como defensa contra el puño dado por el humano, y luego da un puño izquierdo. En la secuencia 6-10 el robot se protege con la pierna izquierda de la patada derecha dada por su oponente.



**Figura 65. Capturas virtuales de la reacción del robot a estímulos.**

Finalmente los autores prueban estas interacciones conectando la salida del sistema virtual al hardware de robot humanoide UT- $\mu$ 2 en donde el movimiento del humano es medido con un sistema óptico de captura de movimiento y luego el modelo genera el movimiento del robot humanoide para la situación actual. La salida del sistema es primero alimentada dentro del interpolador de patrones de movimiento en tiempo real en el cual el hardware toma en cuenta la suavidad y estabilidad del movimiento que luego son enviados al controlador del robot como referencias de ángulos de articulación. En la Figura 66 se observa el resultado experimental de interacción del robot con el humano.

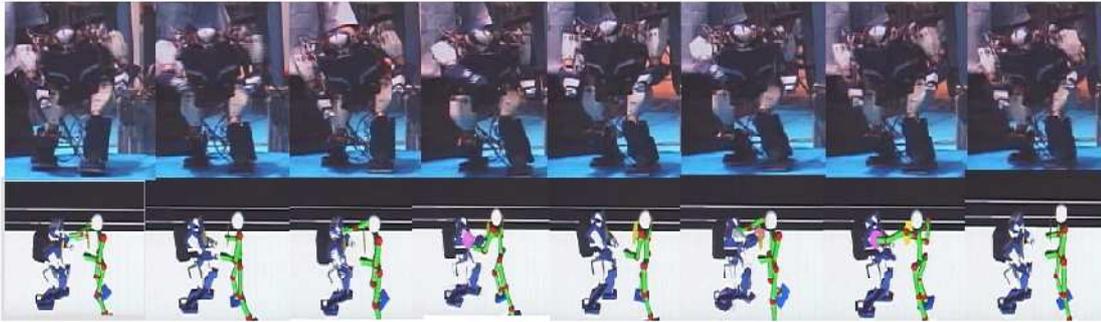


Figura 66. Capturas en tiempo real de la reacción del robot a estímulos.

### RESUMEN APRENDIZAJE Y RECONOCIMIENTO DE PATRONES DE MOVIMIENTO CORPORAL EN SERES HUMANOS.

#### 1. Naturaleza y dificultades del reconocimiento de patrones de movimiento corporal.

Se fundamenta bajo la suposición cualquier tipo de expresión corporal puede ser interpretado como una secuencia de eventos que ocurren en el tiempo. Los seres humano reconocen cualquier tipo de expresión o movimiento corporal incluso antes de que dicha acción sea finalizada, teniendo ventaja en poder ejecutar una acción como reacción a estímulo dado. Extrapolar tal situación a una maquina artificial, requiere de la utilización de modelos que represente las diferentes expresiones y permitir un contexto de interacción hombre maquina. Se requiere además de un proceso de predicción para prever un espacio de posibles reacciones hipotéticas. Las HMM's en este caso modelan la evolución de un patrón de movimiento corporal y los patrones de interacción entre dos seres humanos..

#### 2. Estructura de HMM para reconocimiento de patrones de movimiento.

- **Lo que representa un HMM:** un HMM de bajo nivel modela una secuencia de movimientos de algún tipo de expresión determinada por una persona. Un HMM de alto nivel correlaciona los patrones de movimientos del HMM de bajo nivel como un proceso interactivo entre dos personas.
- **Topología:** Por lo general se utiliza una topología de izquierda derecha debido a que la secuencia de movimientos se entienden como un proceso análogo a la pronunciación a una palabra, donde ciertas posiciones del cuerpo se pueden dar si y solo si han pasado por otras posiciones sin saltarse posiciones intermedias.
- **Numero de estados:** Se determina a prueba y error hasta que un número adecuado de estas represente en forma optima un patrón de movimiento específico.
- **Observaciones O:** Al igual que en aplicaciones anteriores, la selección de las observaciones depende de las variables medidas por sensores utilizados para registrar un patrón de movimiento. Las dimensiones de estas observaciones depende del número de sensores utilizados, o del producto de un proceso de extracción de características que comprima los datos, pues sensar patrones de movimiento requiere de varios sensores. La distribución de estas observaciones sobre los estados depende de las características representativas de las señales a un patrón de movimiento determinado.

#### 3. Entrenamiento de los modelos de reconocimiento de movimiento con HMM's.

El sistema es sometido a prueba con patrones de movimiento predefinidos y supervisados. Igualmente se extrae los patrones de interacción de dos personas por medio del esquema de aprendizaje por demostración para que el sistema correlacione los patrones de movimiento de cada uno.

#### 4. Reconocimiento de patrones de interacción y movimiento.

Se somete el sistema a múltiples situaciones de interacción con un ser humano y en base a las acciones ejecutadas por el ser humano, el sistema planea una estrategia de reacción. Primero identifica el patrón de movimiento del humano, luego establece los posibles movimientos de reacción, ejecuta la acción y verifica que los patrones de interacción sean los adecuados.

### 5. TOOLBOXES Y OTRAS HERRAMIENTAS INFORMÁTICAS EN HMM's

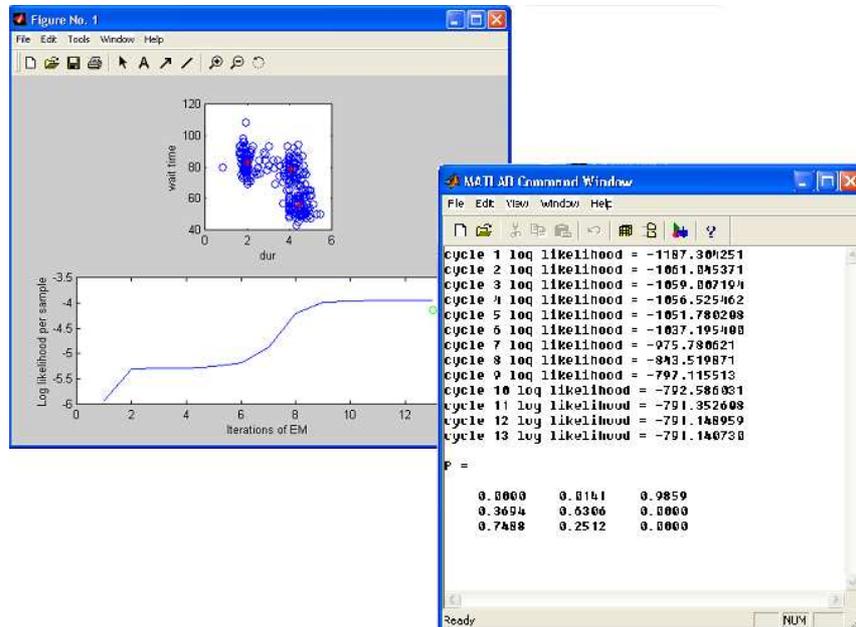
En la actualidad se pueden encontrar en internet varios sitios Web en los cuales se pueden obtener Toolboxes, Software y otros elementos de descarga libre de algoritmos de HMM's. Se menciona brevemente algunas características de estas herramientas:

#### 5.1. Dhmm.tar y Hmm.tar

- **Sitio web:** [www.gatsby.ucl.ac.uk/~zoubin/software.html](http://www.gatsby.ucl.ac.uk/~zoubin/software.html)
- **Desarrollado por:** Zoubin Ghahramani.

Posee un toolbox de MATLAB muy sencillo y liviano donde muestra el funcionamiento de una HMM discreto y un HMM continuo de una sola mixtura respectivamente. Se puede considerar este toolbox como un buen elemento para comenzar a interactuar con HMM's. En la página de descarga se indican los procedimientos de instalación y ejecución. En la Figura 67 se muestra el modelamiento de una secuencia de datos de 2 variables en donde las medias de las funciones gaussianas se centran en los datos dados. Adicional se muestra una grafica de probabilidad con respecto a cada iteración y la matriz de transición (P, según el programa) obtenido después de la etapa de entrenamiento.

Figura 67. Ejecución del HMM.tar

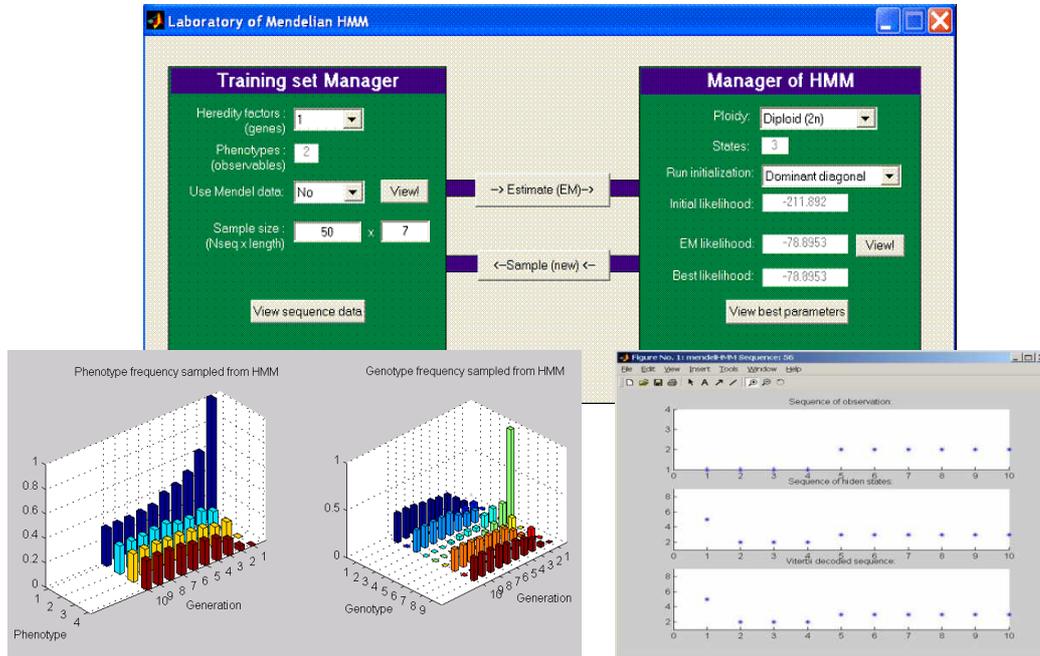


### 5.2. MendelHMM

- **Sitio web:** <http://www.math.uit.no/bi/hmm/>
- **Desarrollado por:** Steinar Thorvaldsen.

Es un tutorial toolbox de uso académico libre para MATLAB. Presenta una Interfaz grafica que manipulación de parametros del programa. El software trata básicamente de modelar la secuencia de caracteres dominantes o recesivos de genes de diferentes especies de guisantes como se conoce en el experimento de Gregor Mendel (1822-1884). EL toolbox permite manipulación de los datos de entrada y de los parámetros del modelo HMM. En la Figura 68 se observa algunas imágenes de la Interfaz gráfica y la simulación de secuencias de muestras de distribución de observaciones y estados ocultos para diferentes números de fenotipos y genotipos.

Figura 68. Imágenes de interfaz grafica del MendelHMM toolbox.



### 5.3. UMDHMM.

- **Sitio web:** [www.kanungo.com/software/software.html](http://www.kanungo.com/software/software.html)
- **Desarrollado por:** Tapas Kanungo.

Este toolbox en lenguaje C hace la implementación del algoritmo de Forward-Backward, Viterbi, and Baum-Welch. Es un toolbox sencillo donde muestra la utilidad de los algoritmos básicos. Algunos científicos interesados en el toolbox lo han utilizado este para implementarlo en reconocedor de voz y bioinformática.

### 5.4. HTK (Hidden Markov Model Toolkit).

- **Sitio web:** <http://htk.eng.cam.ac.uk/>
- **Desarrollado por:** (CUED) Cambridge University Engineering Department.

Existen otros toolboxes que corren sobre plataformas Linux. Entre ellos está el HTK (Hidden Markov Model Toolkit) utilizado básicamente para reconocimiento de voz y reconocimiento de caracteres, y de secuencias de ADN. La herramienta provee utilidades para el reconocimiento de voz, entrenamiento de HMM, prueba y análisis de resultados. El software soporta HMMs que utilizan densidades de mixturas gaussianas y distribuciones discretas. Las librerías de este toolkit están disponibles en lenguaje C.

### 5.5. HMM\_mat y HMMall

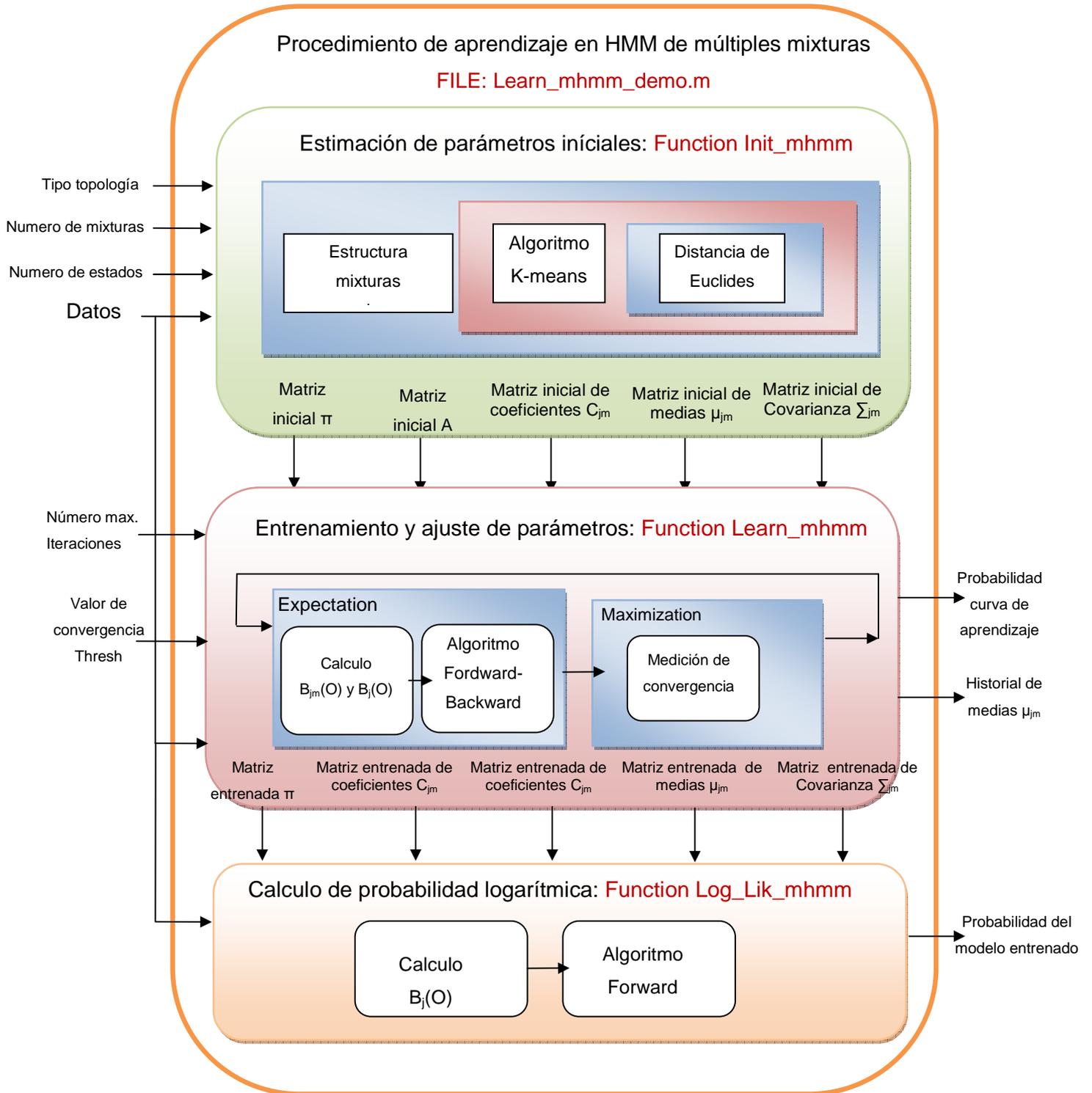
- **Sitio web:** <http://stardec.ascc.neu.edu/~bba/CBIO3580/HW7.html>
- **Desarrollado por:** Kevin Murphy.

Este toolbox de MATLAB proporciona inferencia y aprendizaje de HMMs con salidas discretas (dhmm's), funciones gaussianas (ghmm's), o mixturas de funciones Gaussianas (mhmm's). Las funciones gaussianas pueden contener matrices de covarianzas de completas, diagonales o esféricas. Soporta además entradas discretas, como en POMDP (Partially Observable Markov Decision Process). A pesar de no poseer una interfaz gráfica, es bastante completo en muchos elementos algorítmicos para el análisis de cualquier secuencia de datos, de cualquier dimensión y tipo.

Este toolbox es uno de los más completos pues tiene algoritmos para modelar eventos discretos y continuos de diferentes mixturas. Entre los archivos ejecutables de aprendizaje de modelos se encuentra: (learn\_mhmm.demo.m) para el aprendizaje de HMM's de observaciones continuas de múltiples mixturas; (learn\_gmm.demo.m) para el aprendizaje de HMM's de observaciones de una sola mixtura. Este programa sería un caso especial del programa (learn\_mhmm.demo.m) donde  $M=1$ ; (learn\_dmm.demo.m) para entrenar HMM's de eventos discretos.

En la Figura 69 se muestra la estructura del programa para el aprendizaje de HMM's de múltiples mixturas en observaciones continuas. Básicamente consiste en 3 bloques de proceso donde se hace una selección de los parámetros iniciales de acuerdo a los datos recibidos, luego aplica el método de entrenamiento por los procedimientos de EM (Expectation- Maximization) de manera iterativa, y por último evalúa el desempeño del entrenamiento.

Figura 69. Estructura del programa Learn\_mhmm\_demo.m



### 6. CONCLUSIONES

En este documento se hizo una breve contextualización de las Cadenas Ocultas de Markov como maquina de aprendizaje dentro de la Inteligencia artificial. Se le describió además como método estadístico para modelar procesos que se desarrollan en el tiempo, teniendo en cuenta la estructura básica de esta máquina, sus elementos, su mecanismo de aprendizaje y las tareas que puede ejecutar.

Se observó la robustez de esta máquina con la descripción de 4 aplicaciones de diferentes naturalezas corroborando la flexibilidad de este método para implementarse a diferentes situaciones que se ajusten a sus bases teóricas.

Particularmente, y como objetivo principal del documento, se comprobó como naturaleza de cada aplicación era compatible con la teoría de HMM. Se evidenció además, que la implementación de tales aplicaciones se encuentra ante situaciones de común denominador. Entre estas situaciones se resalta que la selección de la estructura (topología, número de estados) requiere de un conocimiento considerable por parte del diseñador sobre la naturaleza de señales a modelar. En algunos casos se requiere utilizar métodos previos de aprendizaje no supervisado para identificar clases dentro de la información. Análogamente, la selección del tipo de observaciones, su dimensión, y las funciones de densidad probabilística, también están ligados a naturaleza de la aplicación, al número de variables requeridas para monitorear el proceso y si es necesario, al proceso de pre-procesamiento de los datos para la obtención de vectores característicos.

Por último, se utilizó algunos algoritmos para el entrenamiento, evaluación y prueba de HMM disponibles en la WEB. Se hizo una descripción genérica de las características de tales herramientas y se analizó, como un caso general, el procedimiento de entrenamiento de una HMM de múltiples mixturas de funciones gaussianas de variable continua y multidimensional en una algoritmo (`learn_mhmm.demo.m`) contenido en el toolbox de MATLAB, `HMM_mat`, seleccionado por ser sencillo y completo en su estructura algorítmica. Existen otros procedimientos descritos por otros algoritmos del mismo Toolbox para el entrenamiento de HMM de una sola mixtura gaussiana (`learn_gmm.demo.m`), o de variables discretas (`learn_dmm.demo.m`) que utilizan los mismo principios del procedimientos para el de múltiples mixturas, considerándose a tales algoritmo como casos específicos.

### BIBLIOGRAFIA

- [1] ABDULLA, Waleed, KASABOV, Nikola. *The concepts of Hidden Markov Model in Speech Recognition*. New Zealand, 1999. Technical Report. University of Otago. Information Science Department.
- [2] AYCARD, Olivier; MARI, Jean- Francois; WASHINGTON, Richard. *Learning to automatically detect features for mobile robots using second-order Hidden Markov Models*. En: International Journal of Advanced Robotic Systems. Vol. 1, No. 4. (2004). p.233-250.
- [3] BERNARDIN, Keni. et. al. *A Hidden Markov Model Based Sensor Fusion Approach for Recognizing Continuous Human Grasping Sequences*. Karlsruhe, Alemania. University of Karlsruhe. 13P.
- [4] BOESNACH, Ingo. et. al. *Classification of phases in human motion by Neural Networks and Hidden Markov Models*. Karlsruhe, Germany. University of Karlsruhe. Institute for Algorithms and Cognitive Systems. 6P.
- [5] BURLET, Julien; FRAICHARD, Thierry; AYCARD, Olivier. *Robust Motion Planning using Markov Decision Processes and Quadtree Decomposition*. En: International Conference on Robotics & Automation: Proceeding of the 2004 IEEE. (2004: New Orleans, LA, EEUU).
- [6] BURLET, Julien; FRAICHARD, Thierry; AYCARD, Olivier. *Robust navigation using Markov Models*. En: International Conference on Intelligent Robots and Systems: Proceedings of IEEE/RSJ. (August 2-6, 2005, Edmonton, Alberta, Canadá).
- [7] DIETTERICH, Thomas G. *Machine Learning for Sequential Data: A Review*. Oregon, USA. Oregon State University. 15P.
- [8] GALLARDO, Domingo et al. *Control local de robots móviles basado en métodos estadísticos y algoritmo genéticos*. España, Universidad de Alicante. Departamento de Tecnología Informática y computación. Grupo i3a: Informática Industrial e Inteligencia Artificial.
- [9] LAUGIER, C. et al. *Geometric and Bayesian Models for Safe Navigation in Dynamic Environments*.

- [10] LIU, Nianjun, LOVELL, Brian. *Gesture Classification using Hidden Markov Models and Viterbi Path Counting*. En: Proceedings VIIth Digital Image Computing: Techniques and Applications. (10-12 Dic. 2003, Sydney). p.273- 282.
- [11] MANTYLA, Vesa-Matti. *Discrete hidden Markov Models with application to isolated user-dependent hand gesture*. Technical Research Centre of Finland : VTT publications 449. 107P.
- [12] MARKOV, Zdravko. *Machine Learning Course*. May 21,1998. University of Sofia. Faculty of Mathematics and Informatics. 15P.
- [13] MITCHELL, Tom. *Machine Learning*. s.l.: McGraw Hill,1997.p.703.
- [14] OGAWARA, Koichi. et al. *Modeling Manipulation Interactions by Hidden Markov Models*. En: Proceedings of the 2002 IEEE/RSJ: Intl. Conference on Intelligent Robots and Systems. (October 2002, Lausanne, Switzerland). p.1096-1101.
- [15] OROPEZA Rodríguez, José Luis. *Algoritmos y métodos para el reconocimiento de voz en español mediante sílabas*. En: Computación y Sistemas, Vol.9, No.3.p.270-286.
- [16] PARDAS, Montse, BONAFANTE, Antonio. *Facial Animation parameter extraction and expression recognition using Hidden Markov Models*. Barcelona, España. Universitat Politecnica de Catalunya. 30P.
- [17] RABINER, Lawrence R. *A tutorial on Hidden Markov Models and Selected Applications in Speech Recognition*. En: Proceeding of IEEE. Vol. 6, No. 2. (Feb 1989).p. 257-286.
- [18] RICH, Elaine. *Inteligencia Artificial*. 2 ed. Madrid: McGraw-Hill, 1994. 703p. ISBN 84-841-1858-8.
- [19] RIIS, Soren Kamarik, *Hidden Markov Model and Neural Networks for Speech Recognition*, 1998. Ph.D. Thesis.
- [20] ROWEIS, Sam. *SCIA 2003 Tutorial: Hidden Markov Models*. Toronto, June 29 2003, University of Toronto. 10P.

- [21] STEFAN, Schaal. et al. *Computational Approaches to Motor Learning by Imitation*. EN: Philosophical Transaction of the Royal Society of London. University of Southern California. Computer Science & Neuroscience. 18P.
- [22] STIEFMEIER, Thomas. et al. *Combining Motion sensor and ultrasonic hands tracking for continuous activity recognition in a maintenance scenario*. Zurich, Switzerland. Institute for Computer Systems and Networks. 8P.
- [23] SUNDARESAN, Aravind. et. al. *A Hidden Markov Model Based Framework for recognition of humans from gait sequences*. University of Maryland, College Park. Centre for Automation Research. Department of Electrical and Computer Engineering. 4P.
- [24] TAKANO, Wataru. et al. *Primitive Communication based on Motion Recognition and Generation with Hierarchical Mimesis Model*. En: Proceedings of the 2006 IEEE: International Conference on Robotics and Automation. (May 2006, Orlando, Florida). p.3602-3609.
- [25] TAPUS, Adriana; AYCARD, Olivier. *Searching a Target with a Mobile Robot*. Suiza. Swiss Federal Institute of Technology Lausanne. Autonomous Systems Lab.
- [26] THRUN, Sebastian. *Robotic Mapping: A survey*. Pittsburgh, February, 2002. Carnegie Mellon University. School of Computer Science.
- [27] WOLF, Denis F. et al. *Autonomous Terrain Mapping and Classification Using Hidden Markov Models*. En: International Conference on Robotics & Automation: Proceeding of the 2005 IEEE. (2005:Barcelona,Spain).