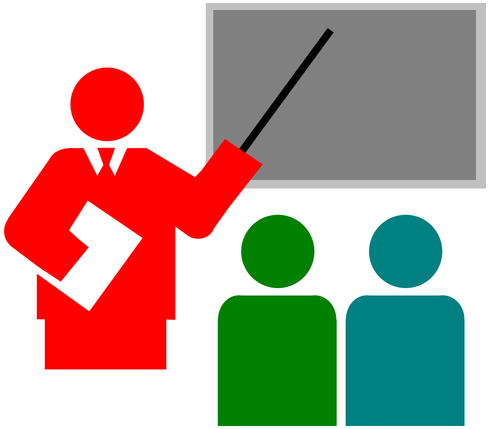
Modelos basados en autómatas probabilísticos

Bioingeniería I

FI-UNER

Organización

- Parte I
  - Autómatas determinísticos
  - Autómatas celulares.
  - Aplicación modelos tejido excitable.



- Parte II
  - Autómatas probabilísticos
  - Modelos de Markov.
  - Aplicaciones

Definición de Autómata

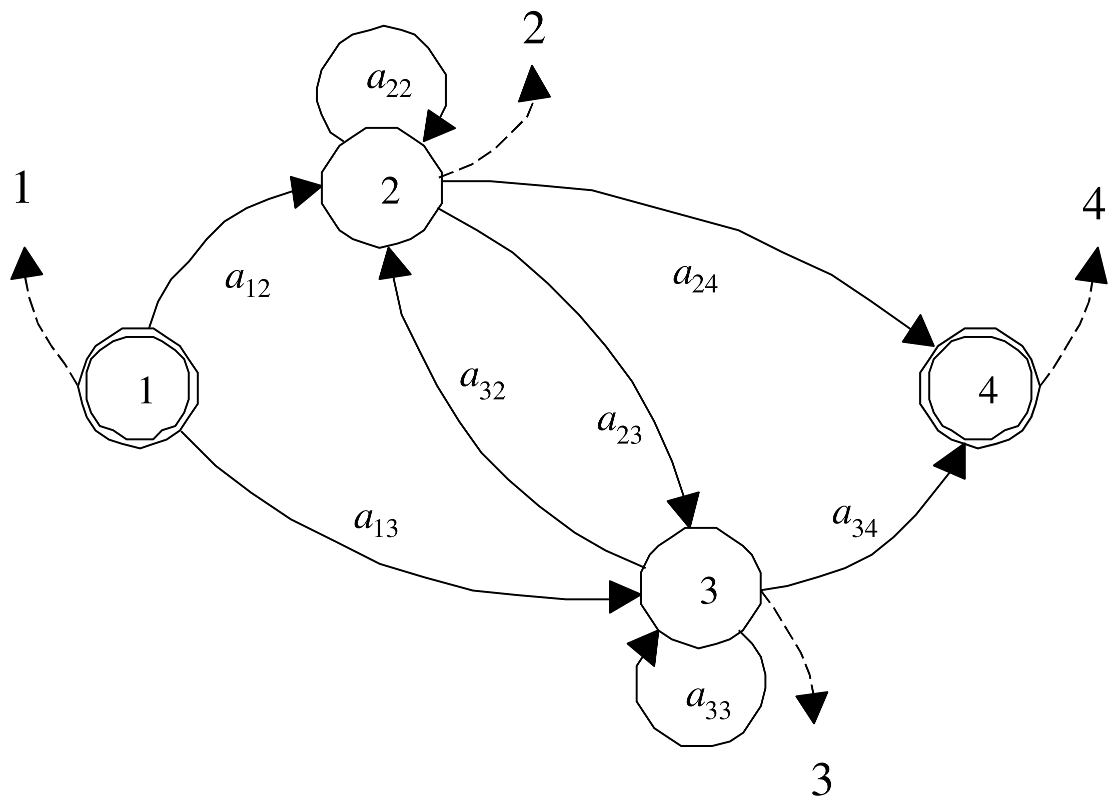
Un Autómata queda especificado por tres conjuntos X, Y y E; y dos funciones δ y β, donde:

- ***X*** **es un conjunto finito de entradas**
- ***Y*** **es un conjunto finito de salidas**
- ***E*** **es el conjunto de estados**
- $\delta : E \times X \rightarrow E$, **la función de** ***transición de estado***
  - si en el tiempo t el sistema está en el estado e y recibe una entrada x, entonces en el tiempo t+1 el sistema estará en el estado $\delta(e,x)$
- $\beta : E \rightarrow Y$, **la función de** ***salida***
  - el estado e siempre da lugar a una salida $\beta(e)$

Autómatas probabilísticos (AEFP)

- A diferencia de los AEF determinísticos:
  - Pueden albergar una descripción probabilística del fenómeno que modelan ⇒ *modelos de señal*
  - 1º nivel de incerteza:
    - En lugar de función de transición de estados se habla de *probabilidades* de transición entre estados.
    - Estados con salida determinística.
  - 2º nivel de incerteza:
    - Cada estado se asocia a uno de los posibles símbolos de un conjunto de salidas mediante su correspondiente *fdp*.

Ejemplo de AEFP



- Las probabilidades de transición desde el estado i al estado j se indican como a_{ij}.
- A cada estado se asocia un símbolo del conjunto finito de salidas.
- En este ejemplo la salida del estado corresponde simplemente con su número.

Modelo de Markov

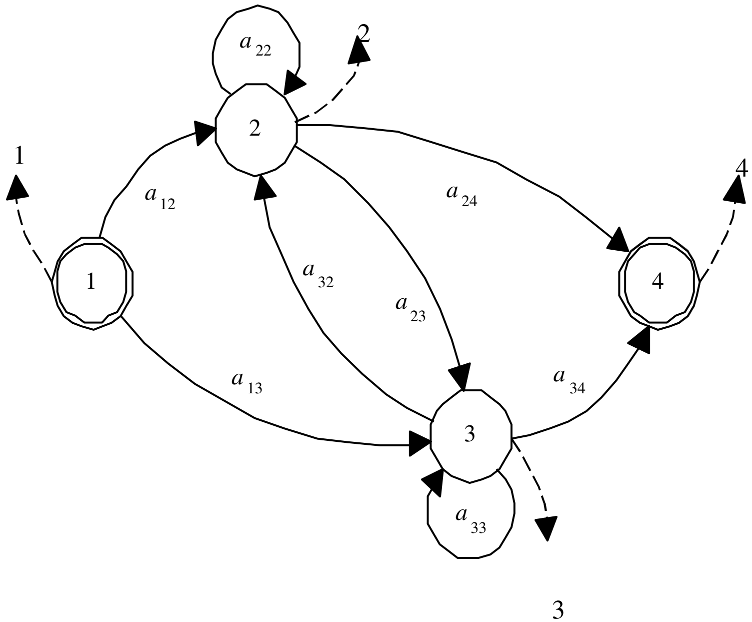
- A este tipo de modelo probabilístico se lo denomina también ***modelo de Markov*** (MM).
- Si el tiempo transcurre a intervalos discretos se trata de un **MM de tiempo discreto**.
- Si las probabilidades de transición sólo dependen de los estados origen y destino, se está en presencia de un proceso de primer orden o **cadena de Markov**.
- Cuando las probabilidades de transición no se modifican con el tiempo también se habla de una cadena de Markov **homogénea**.

Modelo de Markov

- Conjunto de estados: $Q=\{q_1,...,q_n\}$
- Vector de probabilidades iniciales: $\pi=\{\pi_1,...,\pi_n\}$; donde $\pi_i=P(s_0=q_i)$
- Matriz de probabilidad de transiciones: $A=\{aij\}$, donde $a_{ij}=P(s_t=q_j|s_{t-1}=q_i)$
- Vector de observaciones (coincide con el número de estado): $O=\{o_1,...,o_T\}$

MM: $\Lambda = \{A ,\Pi \}$

Ejemplo



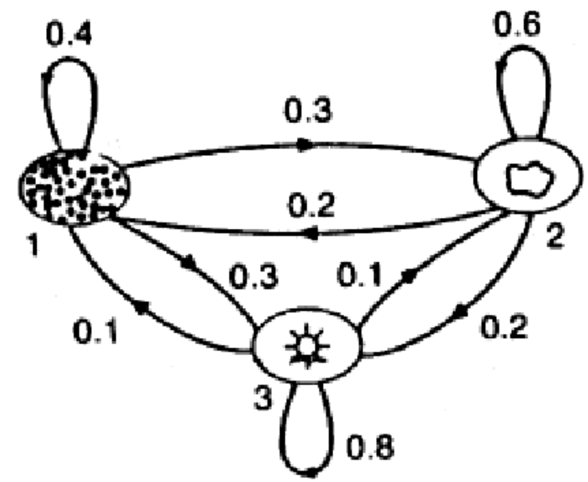
$$A = \{a_{ij}\} = \begin{bmatrix} 0 & 1/2 & 1/2 & 0 \\ 0 & 1/4 & 1/4 & 1/2 \\ 0 & 1/2 & 1/4 & 1/4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

- ¿Qué probabilidad existe de que este modelo genere la secuencia de salida 1, 2, 2, 3, 2, 4?
- La secuencia de estados es la misma: 122324.
- Así, se obtiene la probabilidad total para la secuencia mediante la productoria:

$$p_{122324} = a_{12}a_{22}a_{23}a_{32}a_{24} = \frac{1}{2}\frac{1}{4}\frac{1}{4}\frac{1}{2}\frac{1}{2} = \frac{1}{128}$$

Un ejemplo climático

- Estados:
  1. Lluvia, 2. Nubes, 3. Sol.
- Probabilidad de transición:



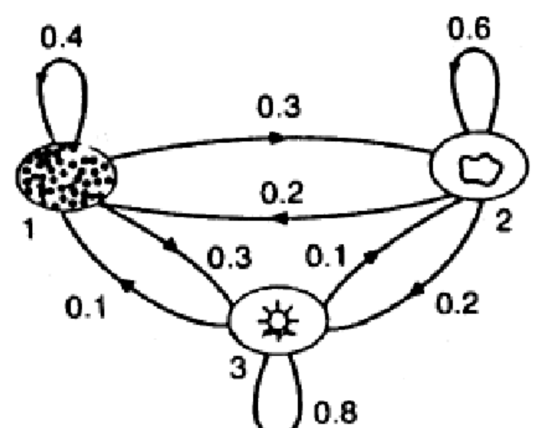
$$A = \{a_{ij}\} = \begin{pmatrix} 0.4 & 0.3 & 0.3 \\ 0.2 & 0.6 & 0.2 \\ 0.1 & 0.1 & 0.8 \end{pmatrix},$$

$$\Pi = \{\pi_i\} = (0.0, 0.0, 1.0).$$

Un ejemplo climático

- Pregunta: ¿con qué probabilidad el tiempo de 8 días será “sol-sol-sol-lluvia-lluvia-sol-nubes-sol”?
- Probabilidad de la observación $O = \{3, 3, 3, 1, 1, 3, 2, 3\}$:

$$P(O \mid Modelo) = P(3, 3, 3, 1, 1, 3, 2, 3 \mid Modelo)$$
$$= P(3) \cdot P(3|3) \cdot P(3|3) \cdot P(1|3) \cdot P(1|1) \cdot P(3|1) \cdot P(2|3) \cdot P(3|2)$$
$$= \pi_3 \cdot a_{33} \cdot a_{33} \cdot a_{13} \cdot a_{11} \cdot a_{31} \cdot a_{23} \cdot a_{32}$$
$$= 1.536 \cdot 10^{-4}$$



Modelos ocultos de Markov

- En cada estado de un MM se emite un determinado símbolo del conjunto de salidas posibles.
- Es por esto que un MM es también conocido bajo la denominación de modelo *observable* de Markov: a partir de la salida se puede "observar" en qué estado se encuentra el modelo.
- Para aumentar su capacidad de modelado, se puede hacer que la función que asocia a cada estado una salida sea una distribución de probabilidades sobre todas las posibles salidas.

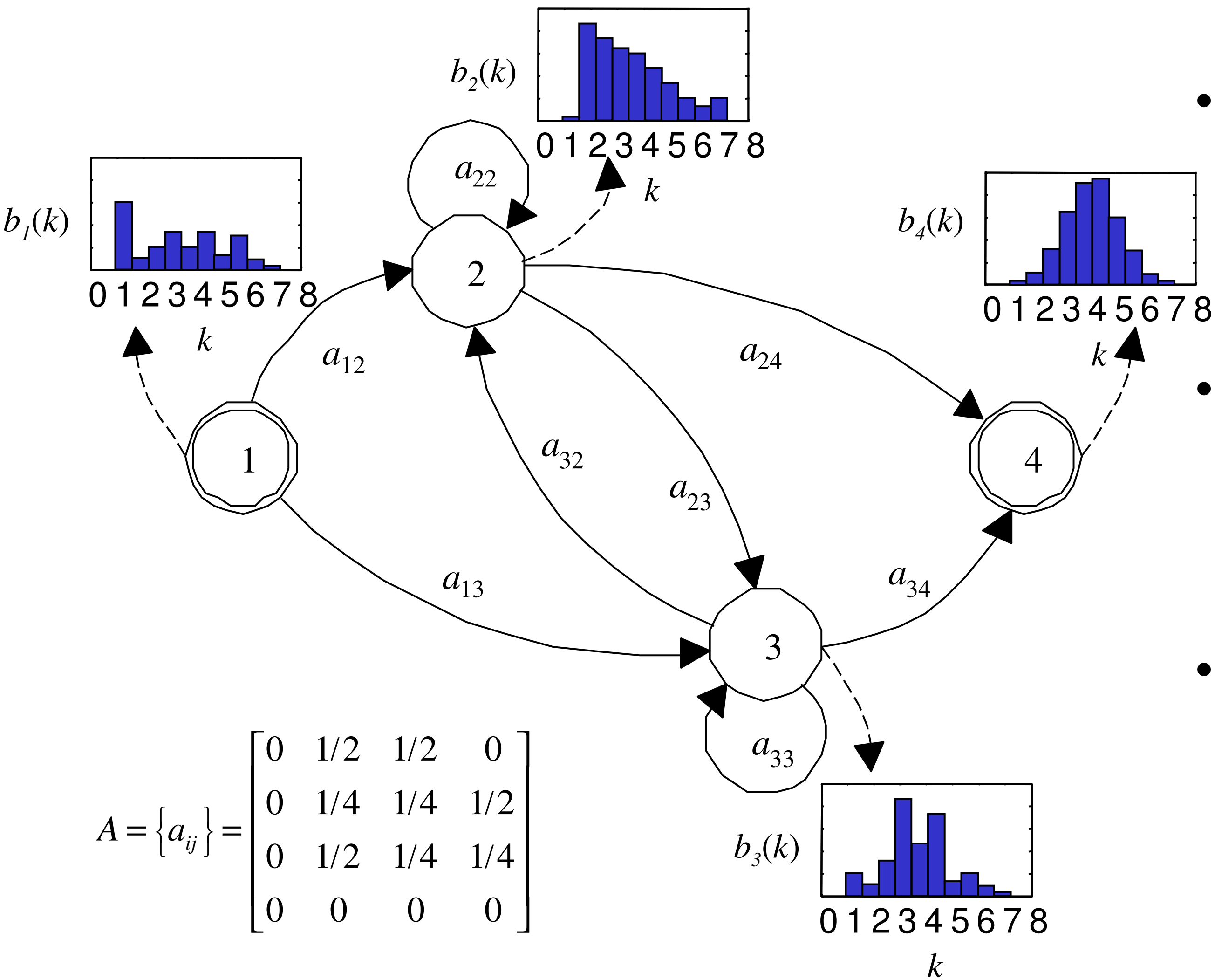
Modelos ocultos de Markov

- Ahora existirá un nuevo parámetro $b_j(k)$ que describe la probabilidad de que en el estado j se observe el símbolo k del conjunto de salidas.
- En estas condiciones nunca se podrá saber con certeza en qué estado está el modelo observando solamente su salida.
- El funcionamiento interno del modelo queda "oculto" y es por eso que se lo denomina modelo *oculto* de Markov (MOM).

Procesos estocásticos a dos niveles

- Los MOM son procesos estocásticos a dos niveles:
  - el de los estados, que no es observable,
  - y el de eventos físicos, que sí es observable.
    - » $B=\{b_j(k)\}$;   j: estados, k: observaciones
- En algunos casos suele hablarse de probabilidades de *emisión* en lugar de probabilidades de *observación*.

Ejemplo: MM $\longrightarrow$ MOM



$$A = \{a_{ij}\} = \begin{bmatrix} 0 & 1/2 & 1/2 & 0 \\ 0 & 1/4 & 1/4 & 1/2 \\ 0 & 1/2 & 1/4 & 1/4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

- Ahora para cada estado existe una *función de distribución de probabilidad* $b_j(k)$.
- En este ejemplo el valor más probable para la salida del estado corresponde a su número.
- Es “posible” que pueda generar algunas secuencias similares al MM.

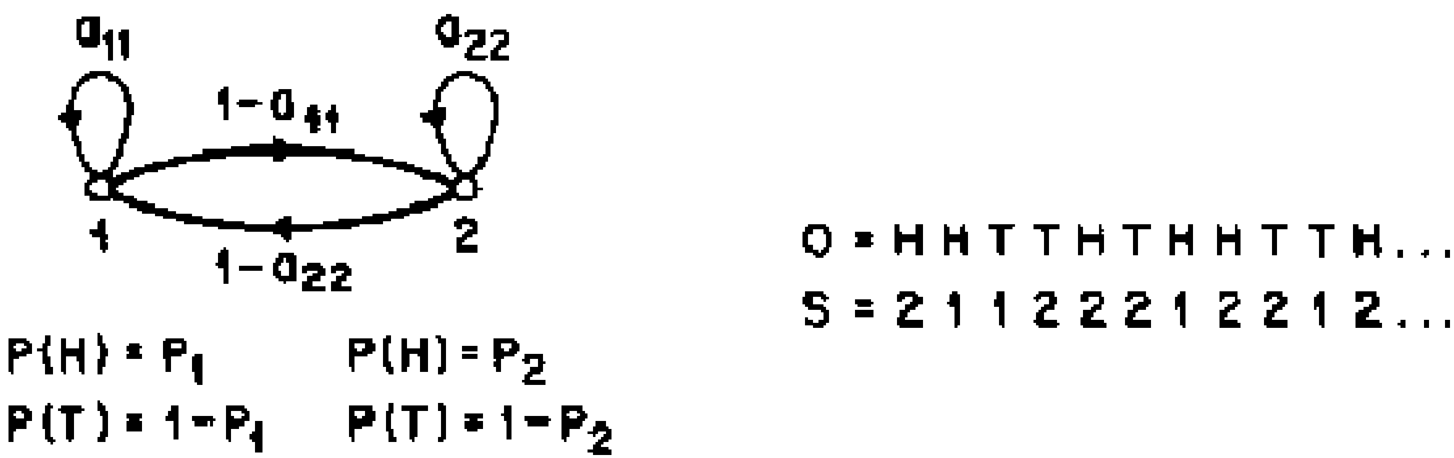
Ejemplo: dos monedas cargadas

- Estamos en una habitación con una cortina que nos impide ver a una persona que tiene dos monedas cargadas de manera diferente.
- Esa persona lanza repetidamente una u otra (aleatoriamente) al aire y nos informa de si sale cara (*H*, por *head*) o cruz (*T*, por *tail*):
  - » H H T H T H H ...

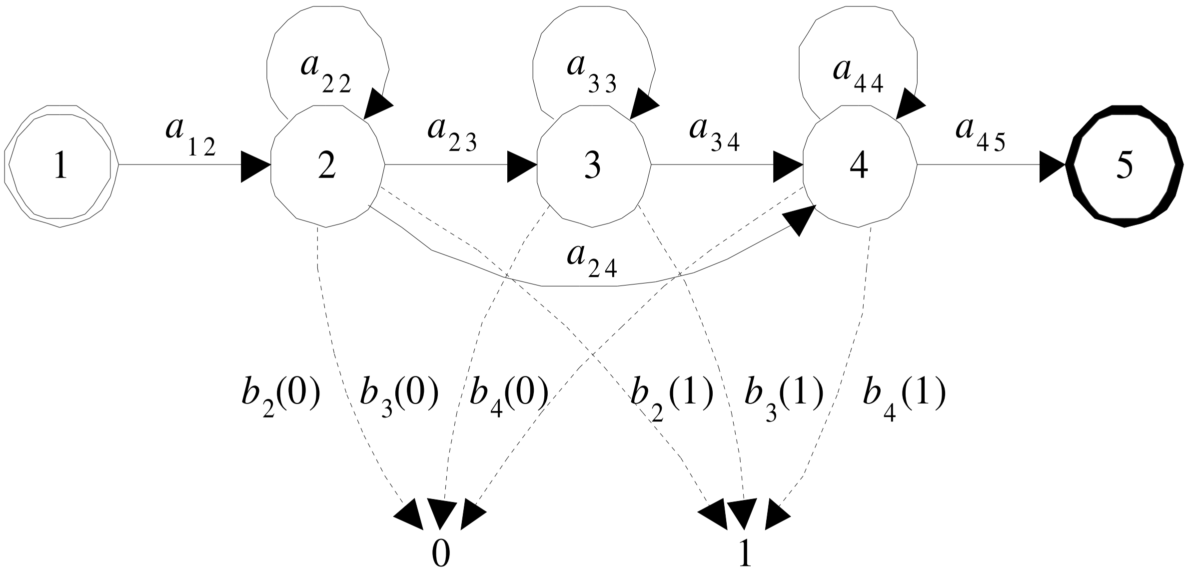
- No tenemos forma de saber cada H o T de qué moneda es:
  - ¿Cómo explicamos/modelamos las secuencias que observamos?

Ejemplo: dos monedas cargadas

- Podríamos modelar este escenario con un MOM:
  - Cada estado correspondería a una moneda.
  - El hecho de que la persona cambia de moneda aleatoriamente correspondería al hecho de cambiar de estado.
  - La asunción markoviana impone que la probabilidad de tomar una u otra moneda dependa exclusivamente de la última moneda lanzada.



MOM Izquierda-Derecha



$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 1/4 & 1/4 & 1/2 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} B = \begin{bmatrix} 0 & 1/3 & 1/5 & 2/3 & 0 \\ 0 & 2/3 & 4/5 & 1/3 & 0 \end{bmatrix}$$

- El estado 1 es el estado inicial y el 5 el final y se denominan no emisores.
- En esta configuración se puede observar la particularidad de que las transiciones se dan solamente de izquierda a derecha.
- Es el más utilizado para el caso de señales.
- ¿Cómo cálculo la probabilidad de la secuencia de salida 0010?

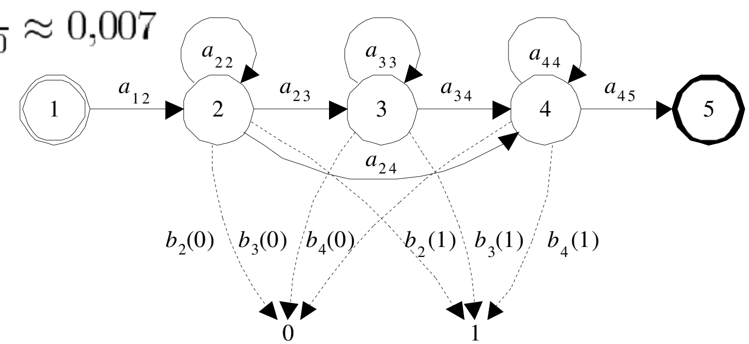
MOM Izda-Dcha

Secuencias de de estados	Probabilidades de transición	Probabilidades de observación	Probabilidades de la secuencia
1, 2, 2, 2, 4, 5	$1\,\frac{1}{4}\frac{1}{4}\frac{1}{2}\frac{1}{2} = \frac{1}{64}$	$\frac{1}{3}\frac{1}{3}\frac{2}{3}\frac{2}{3} = \frac{4}{81}$	$\frac{1}{1296}$
1, 2, 2, 3, 4, 5	$1\,\frac{1}{4}\frac{1}{4}\frac{1}{2}\frac{1}{2} = \frac{1}{64}$	$\frac{1}{3}\frac{1}{3}\frac{4}{5}\frac{2}{3} = \frac{8}{135}$	$\frac{1}{1080}$
1, 2, 2, 4, 4, 5	$1\,\frac{1}{4}\frac{1}{2}\frac{1}{2}\frac{1}{2} = \frac{1}{32}$	$\frac{1}{3}\frac{1}{3}\frac{1}{3}\frac{2}{3} = \frac{2}{81}$	$\frac{1}{1296}$
1, 2, 3, 3, 4, 5	$1\,\frac{1}{4}\frac{1}{2}\frac{1}{2}\frac{1}{2} = \frac{1}{32}$	$\frac{1}{3}\frac{1}{5}\frac{4}{5}\frac{2}{3} = \frac{8}{225}$	$\frac{1}{900}$
1, 2, 3, 4, 4, 5	$1\,\frac{1}{4}\frac{1}{2}\frac{1}{2}\frac{1}{2} = \frac{1}{32}$	$\frac{1}{3}\frac{1}{5}\frac{1}{3}\frac{2}{3} = \frac{2}{135}$	$\frac{1}{2160}$
1, 2, 4, 4, 4, 5	$1\,\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2} = \frac{1}{16}$	$\frac{1}{3}\frac{2}{3}\frac{1}{3}\frac{2}{3} = \frac{4}{81}$	$\frac{1}{324}$

Probabilidad Total $\quad \sum = \frac{77}{10800} \approx 0{,}007$

Probabilidad para todos los caminos posibles para generar la secuencia de salida 0010 en el ejemplo.

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1/4 & 1/4 & 1/2 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad B = \begin{pmatrix} 1/3 & 1/5 & 2/3 & 0 \\ 2/3 & 4/5 & 1/3 & 0 \end{pmatrix}$$



Formalizando...

- Un MOM se caracteriza por:
  - Un conjunto de N estados E.
  - Un espacio de características observables. Un conjunto Y de M símbolos (MOM discretos).
  - Una matriz $A = \{a_{ij}\}$ de probabilidad de transición entre estados. El elemento a_{ij} indica la probabilidad de transitar al estado j si se está en el estado i.
  - Se cumple: $$a_{ij} \geq 0, \qquad 1 \leq i, j \leq N;$$ $$\sum_{j=1}^{N} a_{ij} = 1, \qquad 1 \leq i \leq N.$$

Formalizando...

- Una distribución de probabilidad de emisión de símbolos en cada estado: $B = \{b_j(k)\}$, para $1 \leq j \leq N$, $1 \leq k \leq M$.
- El elemento $b_j(k)$ indica la probabilidad de emitir una observación k en el estado j.
- Se cumple: $b_i(k) \geq 0, \qquad 1 \leq i \leq N, \quad 1 \leq k \leq M;$

$$\sum_{k=1}^{M} b_i(k) = 1, \qquad 1 \leq i \leq N.$$

¡Ojo estados iniciales y finales!

Formalizando...

- Una distribución de probabilidad de estado inicial: $\Pi = \{\pi_i\}$, para $1 \leq i \leq N$.
- El elemento π_i indica la probabilidad de que el primer estado en el que estamos (y desde el que se emite el primer símbolo) sea el estado *i*.
- Se cumple: $\pi_i \geq 0, \qquad 1 \leq i \leq N;$

$$\sum_{i=1}^{N} \pi_i = 1, \qquad 1 \leq i \leq N.$$

- Un modelo queda descrito, pues, con $\lambda = (E, Y, A, B, \Pi)$.

Los problemas básicos de los MOM

- **Problema 1:** El problema de la evaluación. Dada una secuencia de observaciones $O = o_1 o_2 \ldots o_T$ y un modelo $\lambda=(A, B)$, ¿cómo calculamos *eficientemente* la probabilidad de que λ genera a O, es decir, $P(O \mid \lambda)$?
  - Algoritmo *forward-backward*

- **Problema 2:** El problema de la decodificación. Dada una secuencia de observaciones O y un modelo $\lambda=(A, B)$, ¿cómo calculamos la secuencia de estados $x_0, x_1, \ldots x_{T+1}$ que mejor "explica" las observaciones?
  - Algoritmo de Viterbi

- **Problema 3:** El problema del entrenamiento. Dada una secuencia de observaciones O, ¿cómo estimamos los parámetros de $\lambda = (A, B)$ para que $P(O \mid \lambda)$ sea máxima?
  - Algoritmo *EM*

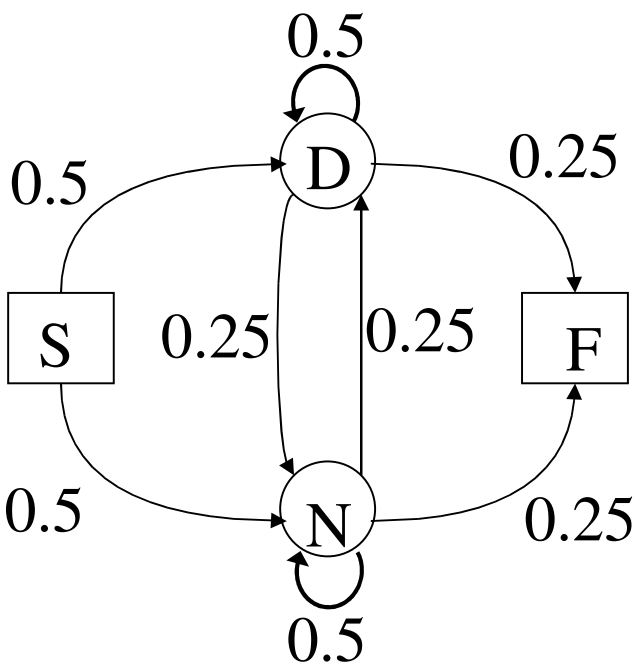
Aplicaciones

Los autómatas como modelos de señales

- Modelos de señales temporales
- Utilización en:
  - Clasificación: reconocimiento de patrones, señales 2D (imágenes) o N-dimensional.
  - Generación de señales sintéticas

Bioinformática

- **Generación de secuencias de aminoácidos:**



Estado D: codifica proteínas

P(A)=P(T)=P(G)=P(C)=0.25

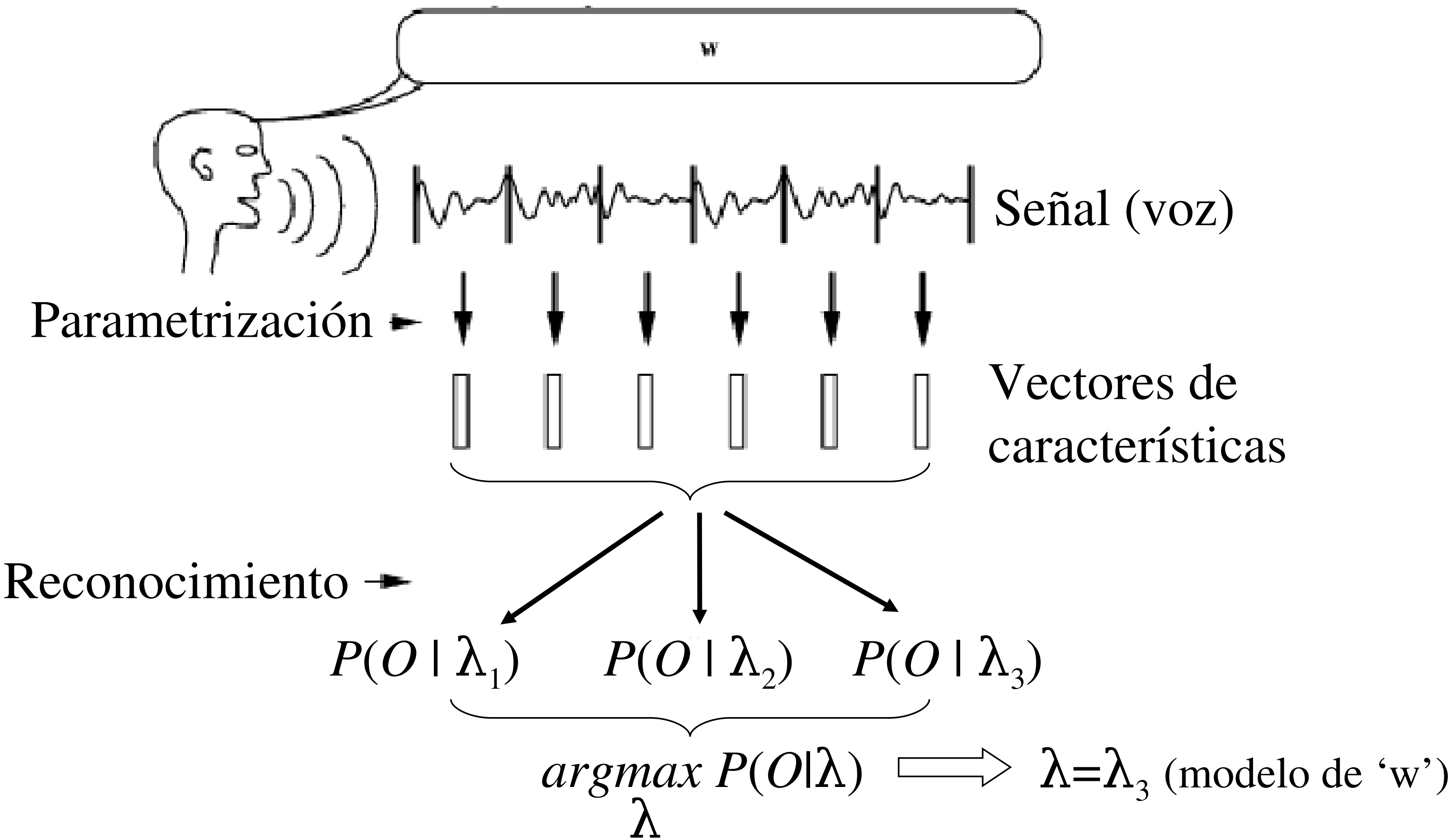
Estado N: no codifica proteínas

P(A)=P(T)=P(C)=0.1; P(G)=0.7

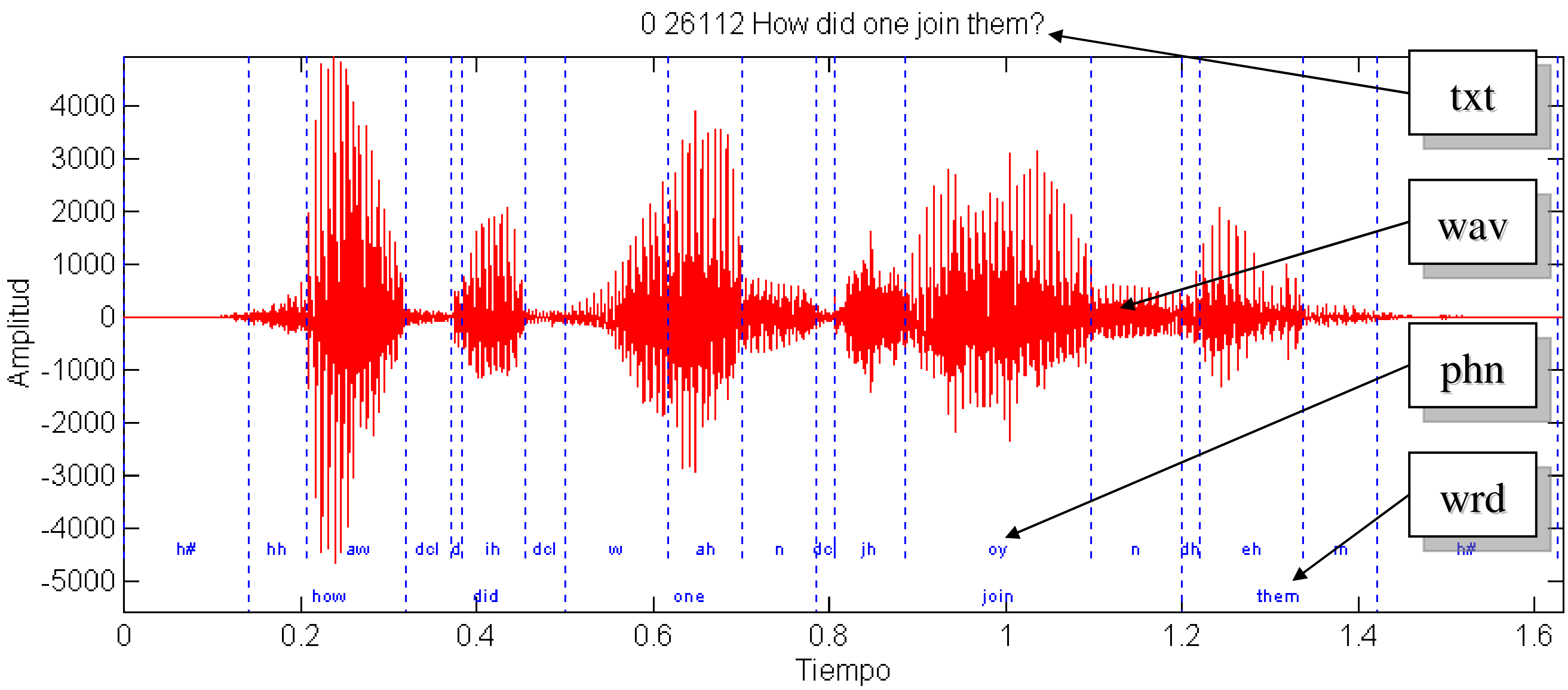
Secuencia observada: ATGTTACTAC…..GGGGTGGAG…

Reconocimiento de Patrones

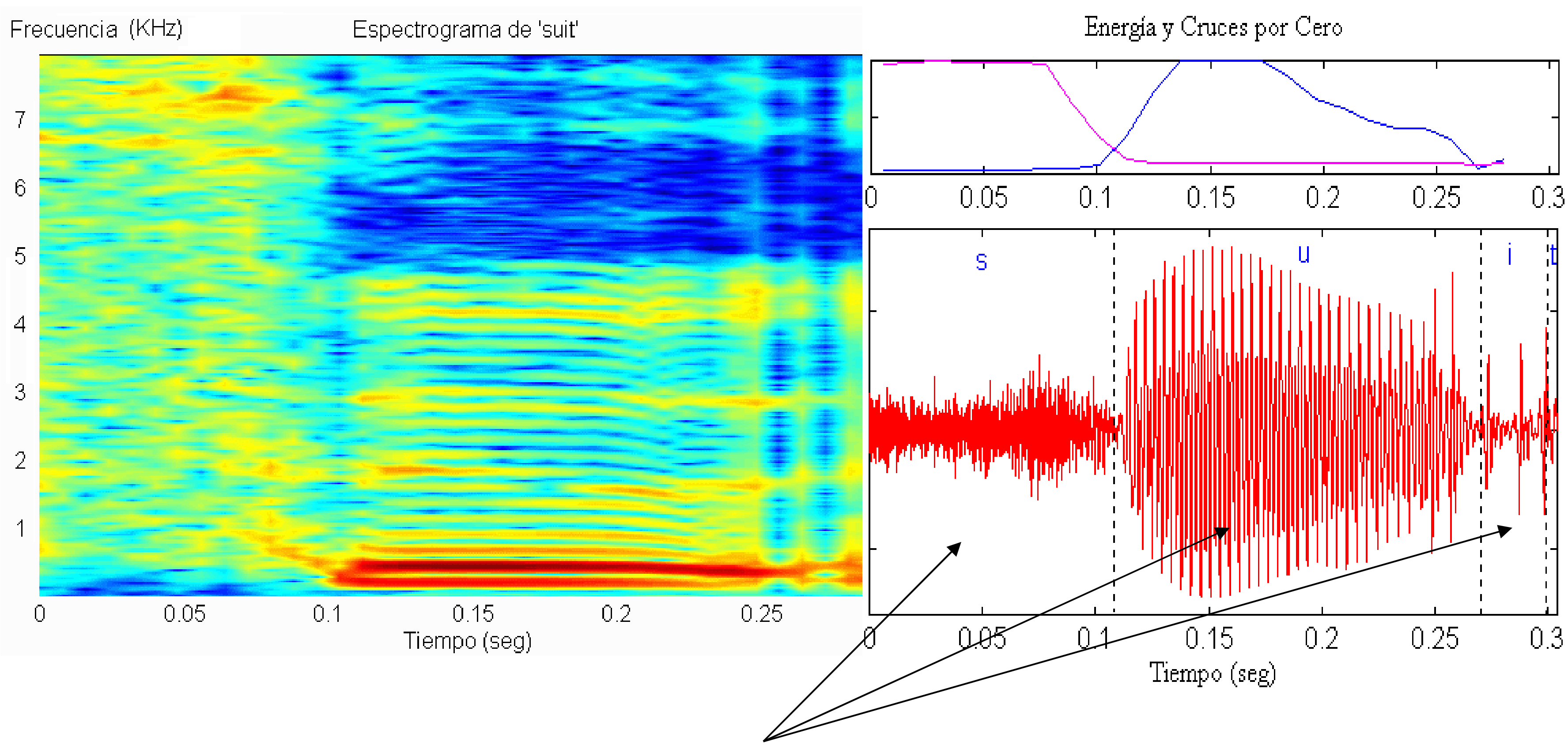
Ejemplo con habla



“Estructura”de la Señal De Voz



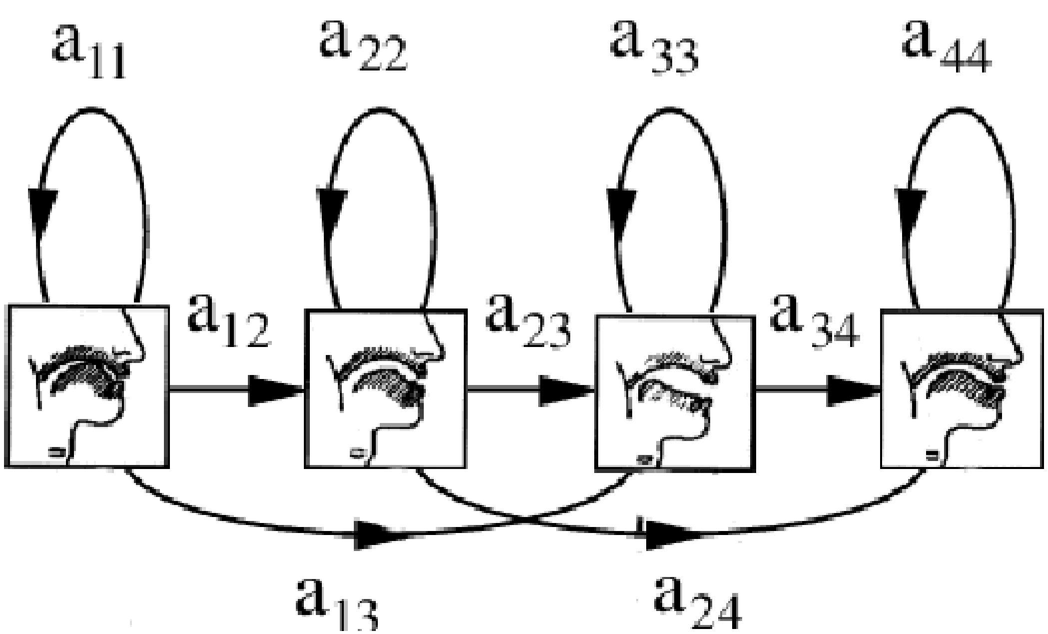
Modelado acústico de fonemas



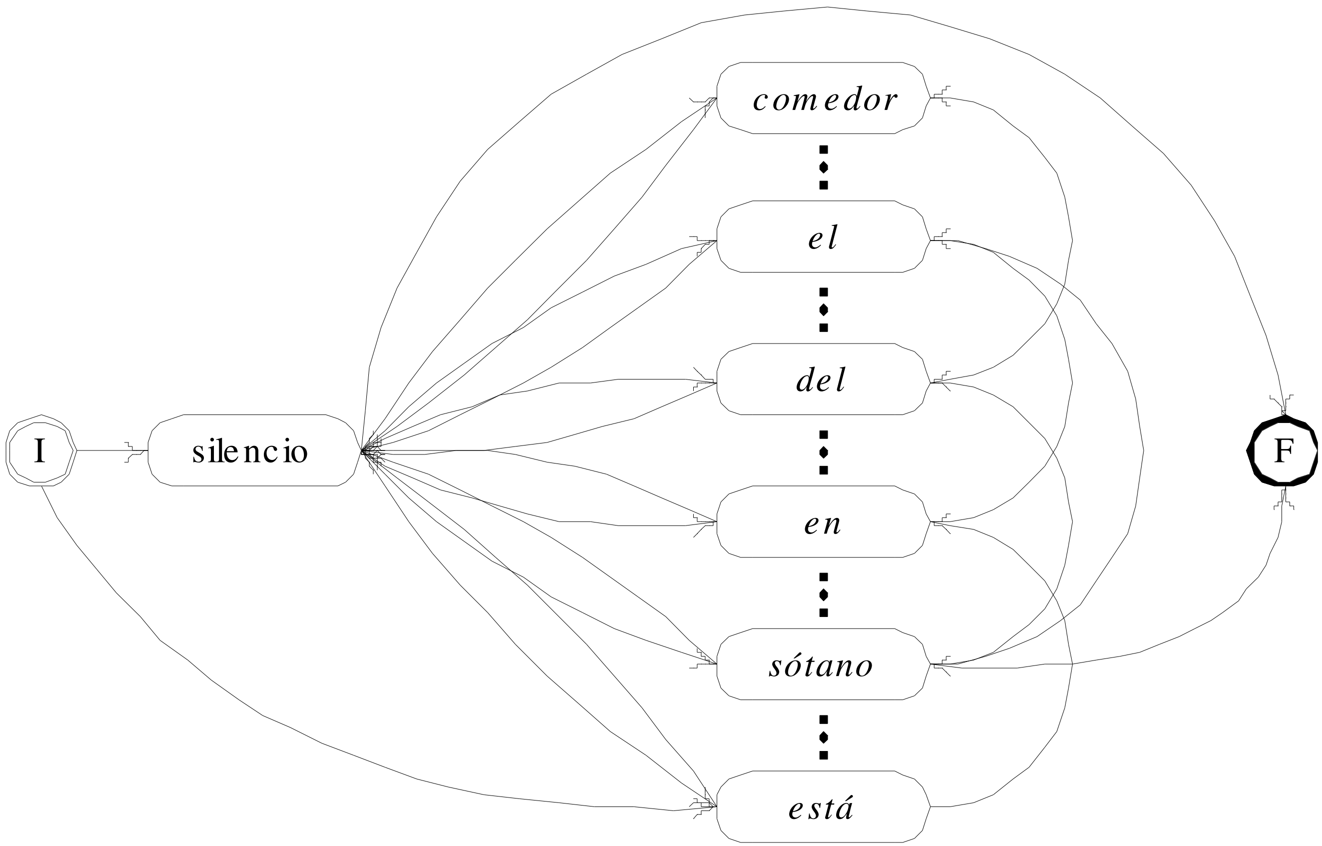
¿Estados?

Modelado acústico de fonemas

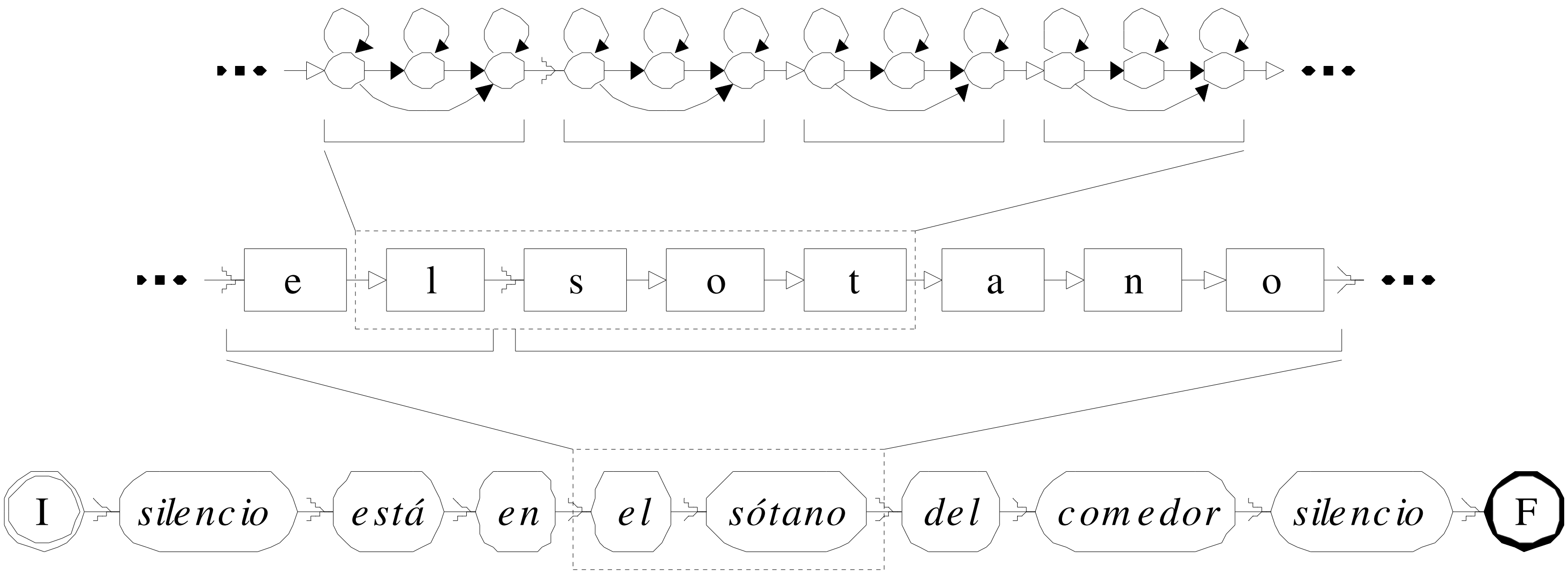
- Asimilamos estados a configuraciones del aparato fonador. De una configuración se puede pasar a otra de acuerdo con ciertas reglas probabilísticas.



Modelado del lenguaje



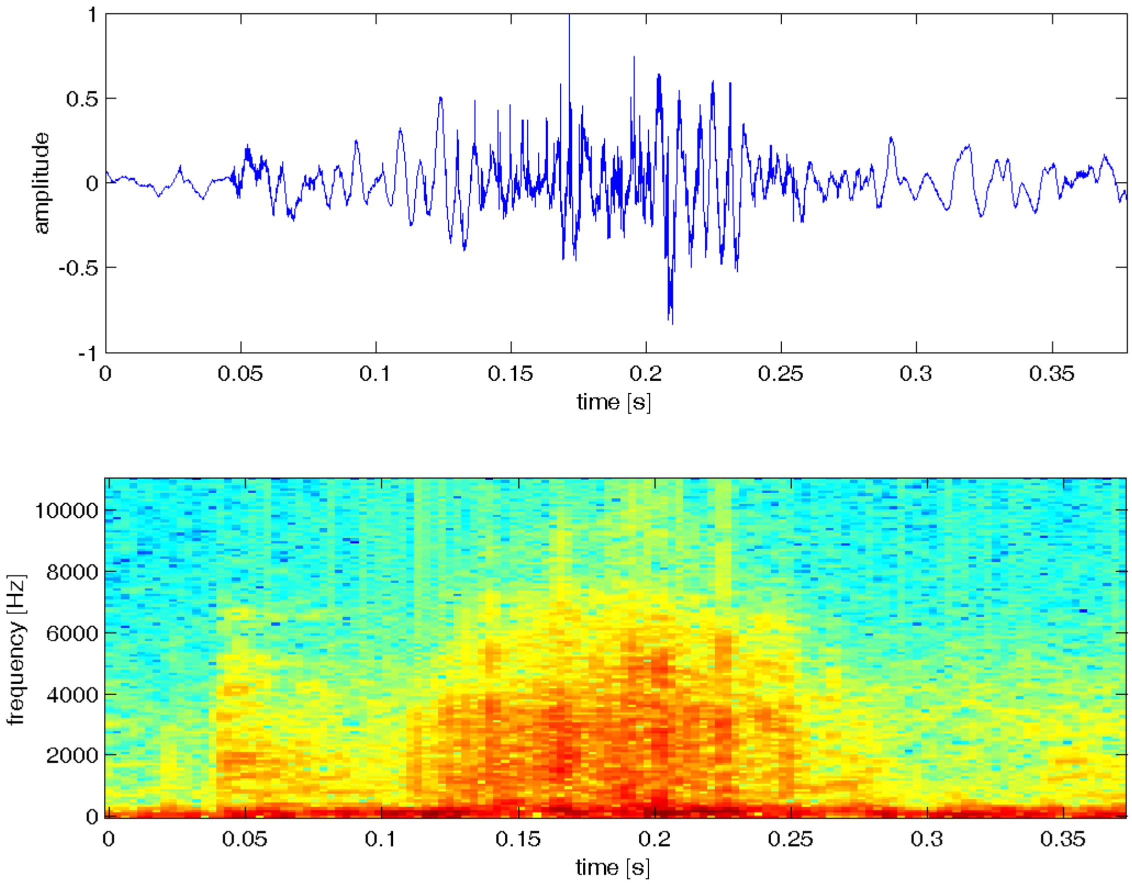
Modelado compuesto



- **Modelo compuesto para la frase: *Está en el sótano del comedor*. Se pueden observar los tres niveles de la composición: los estados del modelo acústico, el diccionario fonético y el modelo de lenguaje. En los modelos acústicos se han eliminado los estados no emisores para simplificar el esquema.**

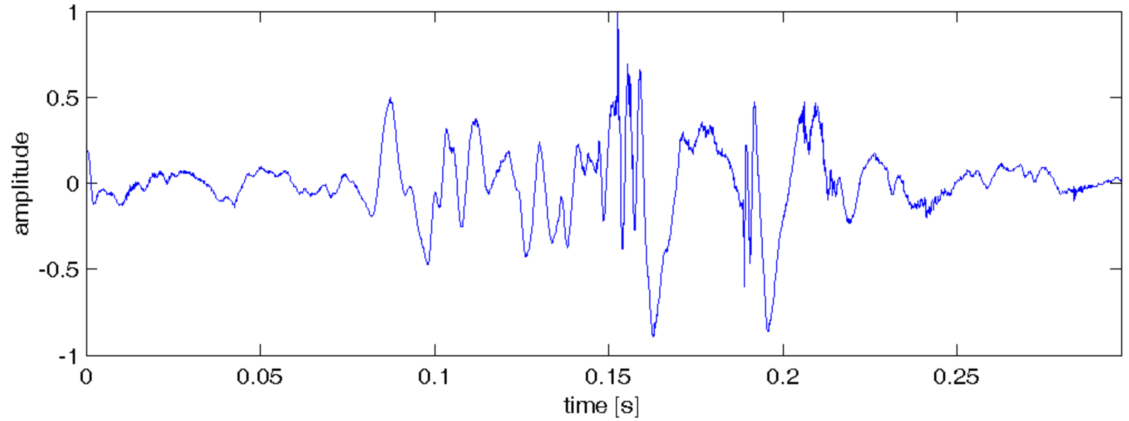
Reconocimiento de sonidos masticatorios en rumiantes

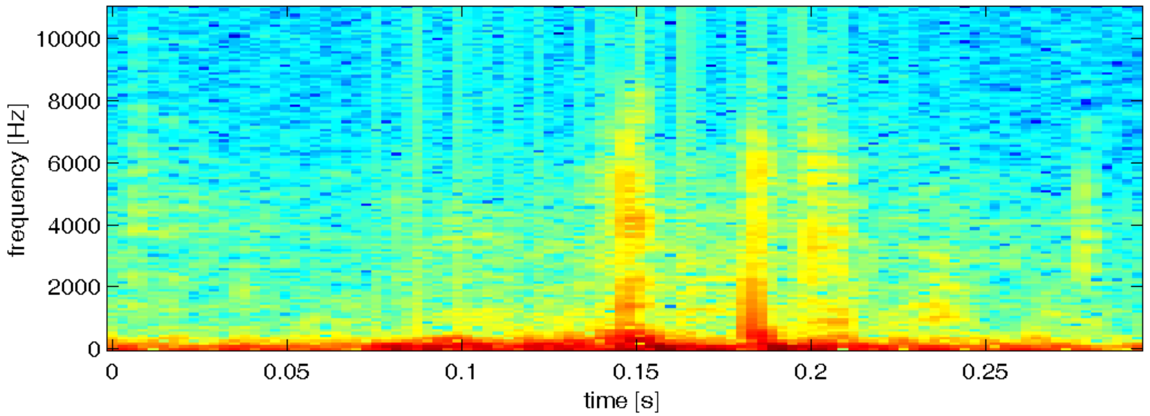
masticación



Reconocimiento de sonidos masticatorios en rumiantes

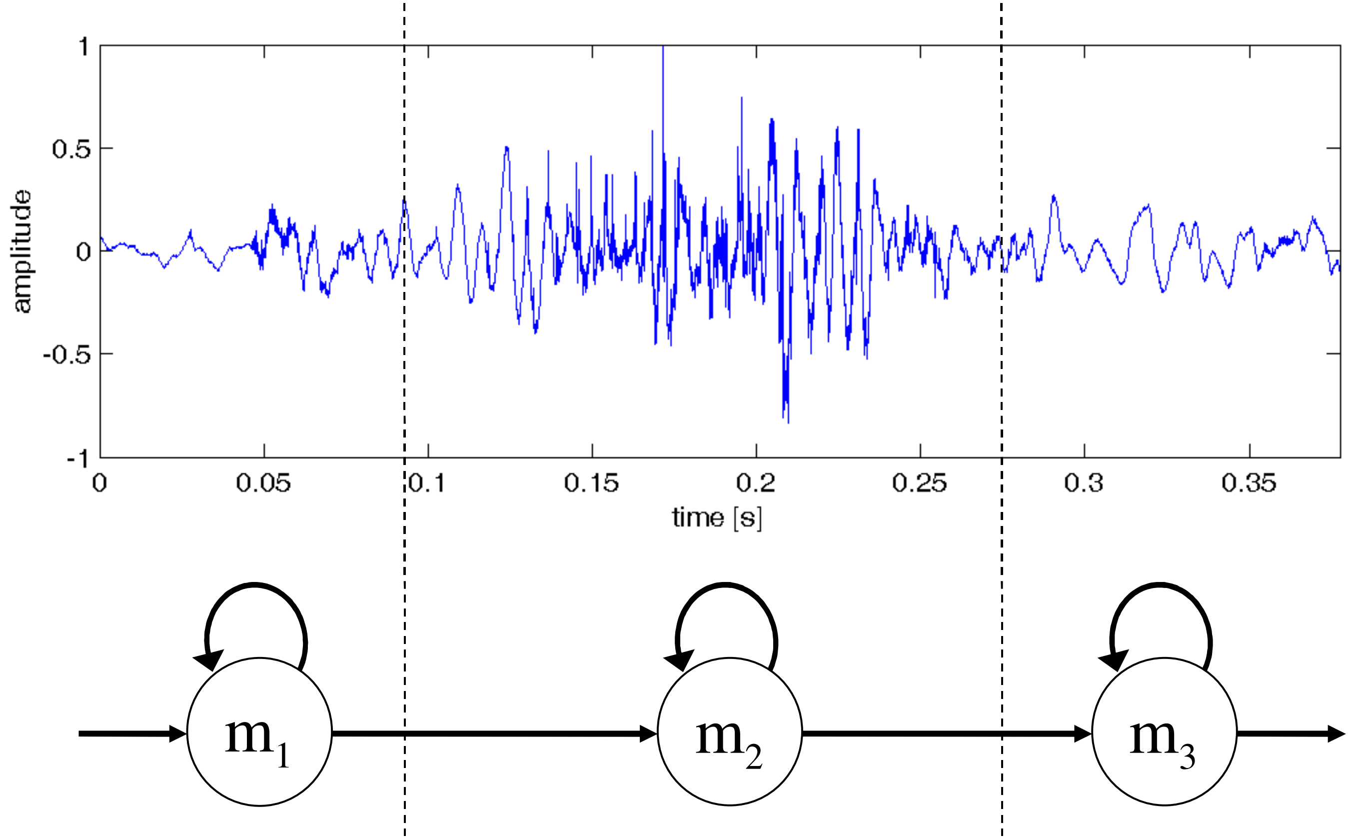
arranque





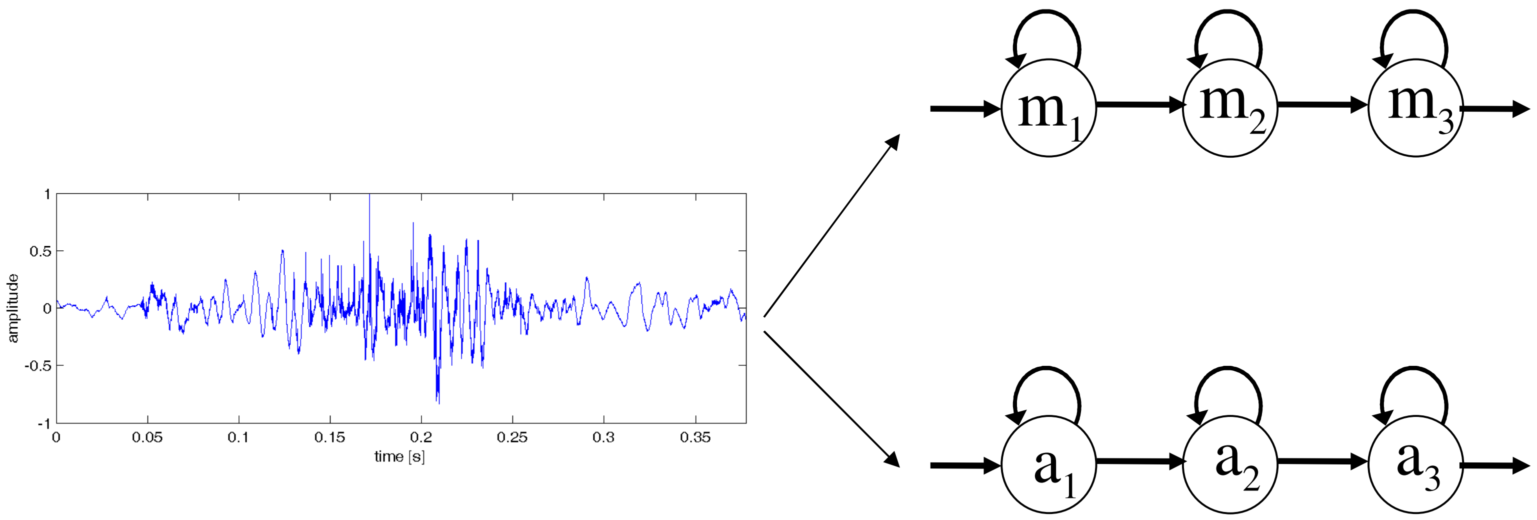
Reconocimiento de sonidos masticatorios en rumiantes

Un autómata por tipo de señal



Reconocimiento de sonidos masticatorios en rumiantes

Decisión por el autómata más representativo



OCR manuscrito

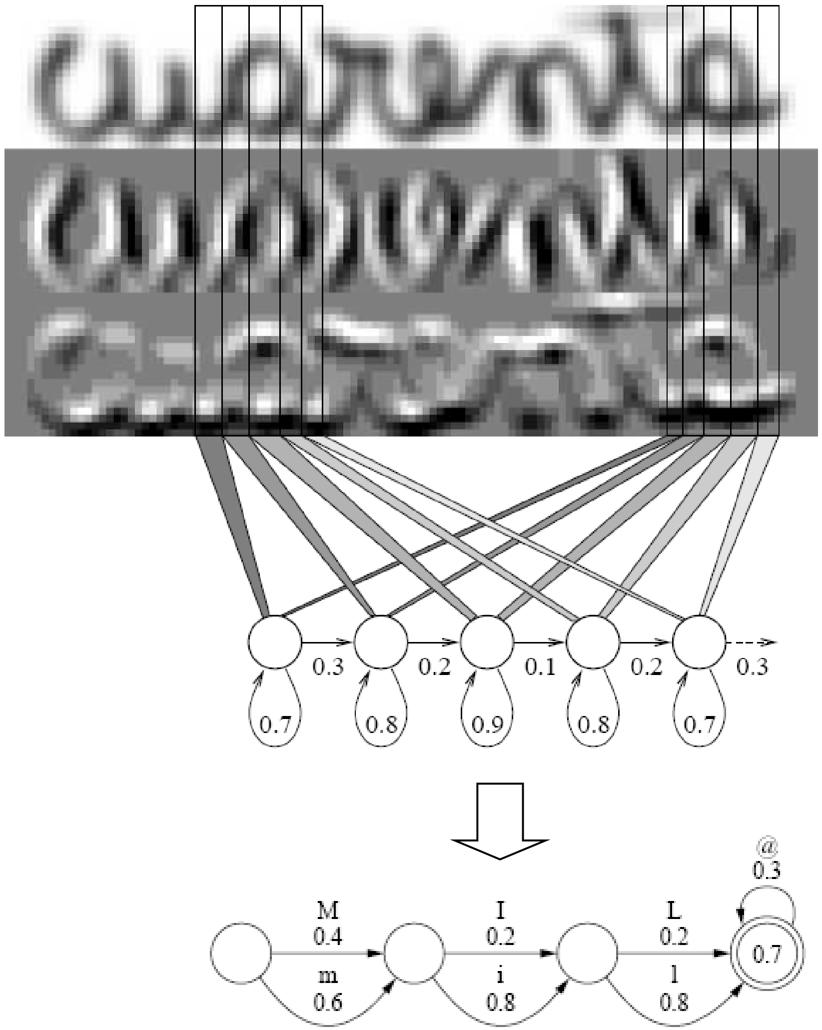
DAR LA OPORTUNIDAD A
QUE UNA PERSONA QUE
LLEVA BASTANTE TIEMPO
CONFIANDO EN TELEFONICA
OFRECIÉNDOSELO OTRO TF. MÓV

LAS HORAS A LAS QUE SE
RECIBE LOS ~~CORREO~~ MENSAJES
SOBRE LOS SERVICIOS DE
MOVISTAR

SIEMPRE LLEGAN TARDE
LAS OFERTAS Y REGALOS

- DEBERÍA TENER UN
SERVICIO DE NOTICIAS
COMPLETAMENTE GRATIS

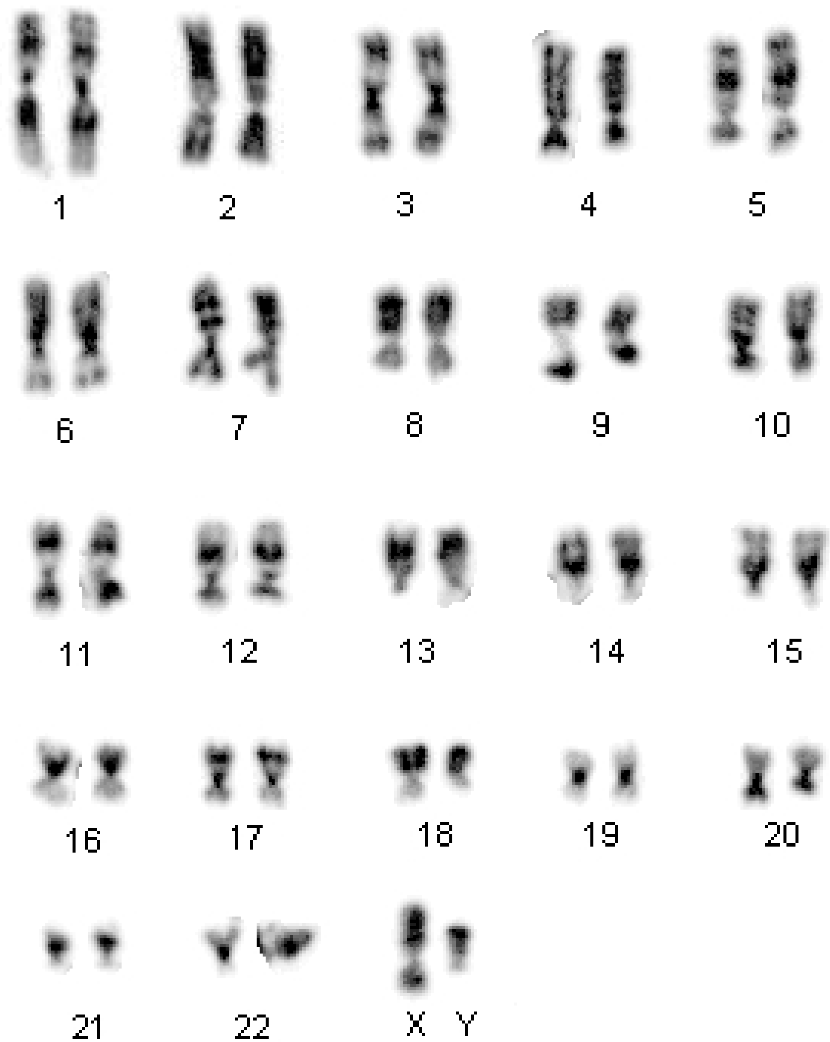
OCR manuscrito

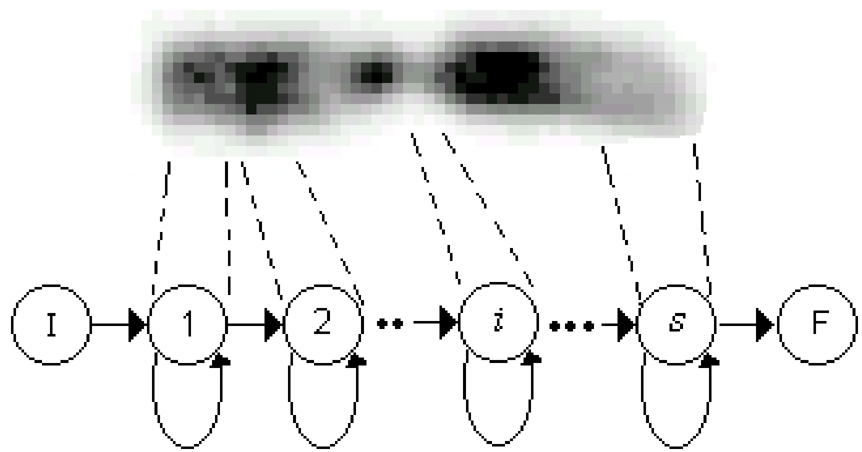


Entrada: imagen

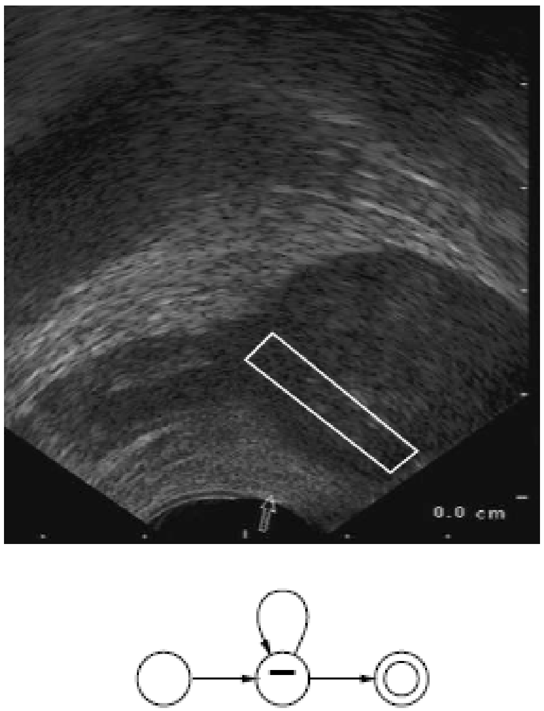
Salida: texto

Clasificación de cromosomas

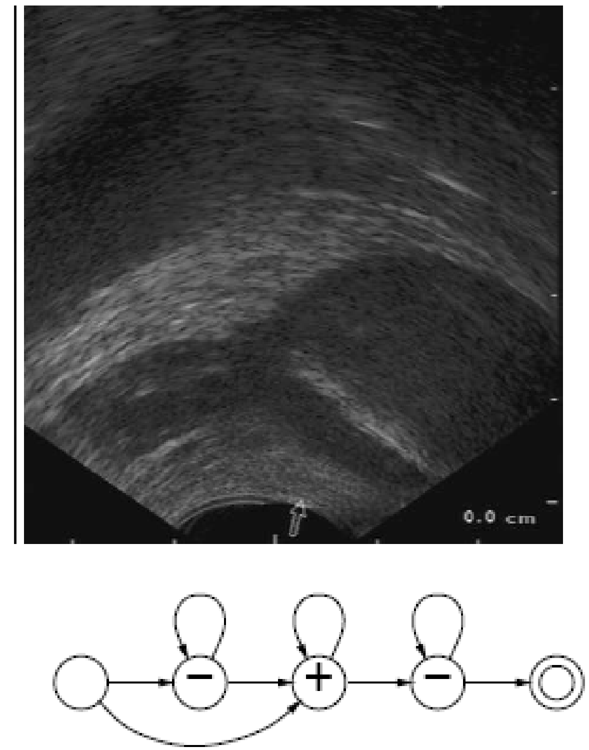




Detección de cáncer de próstata en US



Tejido normal



Tejido con cáncer

Otras señales modelizables

- “Ruidos” cardíacos.
- Señal de VFC
- EEG.
- ECG.
- ....

Bibliografía

- J. Deller, J. Proakis, J. Hansen, “Discrete Time Processing of Speech Signals”. Macmillan Publishing, NewYork, 1993.
- H. Fletcher, “Speech and Hearing in Communication”, Van Nostrand, New York, NY, 1953.
- L. Rocha, “Procesamiento de voz”, I Escola Brasileiro-Argentina de Informática, (Kapelusz, 1987).
- A. M. Borzone, “Manual de Fonética Acustica”, Hachette, 1980.
- Lawrence Rabiner, Biing-Hwang Juang, “Fundamentals of speech recognition”. Prentice Hall. 1993.
- Steve Young et al., The HTK book (http://htk.eng.cam.ac.uk/)
- Frederick Jelinek, “Statistical Methods for Speech Recognition”. The MIT Press. 1998.