

---

---

UNIVERSIDAD CARLOS III DE MADRID

Ingeniería Informática, 3º Curso

Examen de Informática Teórica II

Febrero de 2008

---

---

Duración: 3 horas

**Pregunta 1 ( 2 puntos)**

Dada una Máquina de Turing genérica en la que toda la información que aparece en la cinta está codificada en decimal y el símbolo blanco (es decir, los símbolos de la máquina son 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 y el blanco), indique detalladamente los pasos que habría que realizar para construir una Máquina de Turing equivalente con codificación binaria y el símbolo blanco. Si es necesario la incorporación de estados intermedios o auxiliares, debe indicar todos los estados necesarios.

**Pregunta 2 ( 1,5 puntos)**

Responda razonadamente a las siguientes preguntas:

- ¿Qué es la Máquina de Turing Universal?
- Explique las diferencias entre una Máquina de Turing genérica y la máquina de Turing Universal
- ¿Resuelve la Máquina de Turing Universal el problema de la parada?
- ¿Cuántas máquinas de Turing universales se pueden construir?
- ¿Dónde se encuentran los límites de la computabilidad?

**Pregunta 3 (3,5 puntos)**

Los Himba son un pueblo seminómada del Norte de Namibia que conserva todavía el estilo de vida que tenía hace siglos.



En los últimos años se ha constatado en la población Himba un inusual incremento de personas diabéticas. Desde la Organización Mundial de la Salud (OMS) se ha iniciado un proyecto para ayudar a identificar a la población Himba con mayor riesgo de contraer esta enfermedad. Se está considerando la utilización de siete atributos para realizar el diagnóstico; estos atributos se corresponden con valores enteros y se encuentran todos ellos en un rango establecido. A través de estos atributos se desea establecer si una persona es propensa a contraer la enfermedad de la diabetes o no. Para ello, se ha recolectado la información de 176 individuos diabéticos y 96 que no lo son. Se desea que el error de estimación no sea superior al 5% para que este estudio sea de utilidad. Se pide diseñar un modelo computacional basado en redes de neuronas y se debe indicar:

- Describir la arquitectura neuronal a utilizar: modelo de red neuronal, entradas, salidas, estructura interna.
- Describir, si es necesario, el pre-proceso que deberíamos realizar con la información de entrada y salida.

- c) Describir la interpretación de la salida del modelo neuronal.
- d) Describir el proceso de aprendizaje llevado a cabo para diseñar el modelo neuronal: pasos, selección de patrones, etc.
- e) Supongamos que el error obtenido en el modelo neuronal diseñado es superior al 5%. Explicar los motivos que pueden haber provocado este hecho.
- f) Siguiendo con el supuesto de e), describir que podríamos hacer para obtener un resultado mejor.
- g) Supongamos que el error de entrenamiento es menor que el 5%, pero el error de validación es superior al 5% y crece según avanzamos en el proceso de aprendizaje. Determinar los motivos que han podido provocar esta situación.
- h) Siguiendo con el supuesto de g), ¿qué podríamos hacer para evitarlo?

**Pregunta 4 (3 puntos)**

El ADN es una cadena (molécula) formada por bases nucleotídicas (o letras: A, T, G, C). Supongamos que el ADN de un individuo se compone de una secuencia de 19 bases:

C-T-A-G-T-A-G-T-A-G-G-A-C-A-A-T-G-G-C

La estabilidad de una molécula se mide por el número de Gs y Cs respecto al de As y Ts en una misma secuencia. Así porcentajes superiores al 50% de Gs y Cs son indicativos de alta estabilidad. Por ejemplo, en la secuencia anterior hay 9 Gs o Cs y 10 As o Ts, por lo que molécula es inestable.

Se han realizado estudios sobre la mutación de estas bases entre dos generaciones en dos poblaciones pertenecientes a regiones geográficas distintas (África y Europa).

En un estudio con la población europea se ha comprobado las siguientes cadenas entre generaciones de una misma familia:

Generación 1: C-T-A-G-T-A-G-T-A-G-G-A-C-A-A-T-G-G-C  
 Generación 2: C-A-T-C-G-A-C-T-A-G-G-A-A-A-T-C-G-C  
 Generación 3: C-A-T-A-G-A-C-A-A-G-A-A-A-G-A-T-C-T-C  
 Generación 4: C-G-T-A-G-G-T-A-A-G-A-A-A-C-A-T-C-C-C

Como se puede comprobar en la generación 1, la primera base es una "C" y está se conserva en su hijo (generación 2). Sin embargo, en la segunda base, la generación 1 tenía una "T" y ésta se ha mutado por una "A" en su hijo (segunda generación). Por lo tanto, existe cierta posibilidad de mutación para una base entre dos generaciones. Por otro lado, en la población africana se constataron las siguientes mutaciones:

Generación 1: T-C-A-T-G-A-C-T-C-G-A-C-G-T-A-G-T-C-A  
 Generación 2: T-T-A-A-G-G-C-T-C-G-A-C-G-T-A-A-G-T-A  
 Generación 3: T-G-A-A-C-G-C-T-C-T-A-C-G-T-A-C-G-G-A  
 Generación 3: T-A-A-C-C-G-C-A-C-T-A-C-G-C-C-C-T-G-A

Se pide:

- 1) Definir el Autómata Probabilístico que describa la mutación genérica de una base entre un padre y un hijo
- 2) El ADN de un individuo de Mozambique es:

C-T-A-G-T-A-G-T-A-G-G-A-C-A-A-T-G-G-C

Como se puede comprobar, esta molécula es inestable (baja estabilidad). Razonar mediante el modelo del apartado 1) la posibilidad de que sus nietos posean un ADN de alta estabilidad