Tarea 11

Inteligencia Artificial

1. Entregar un documento PDF con todas tus respuestas teoricas. No se aceptan otro tipo de formato.
2. En esta tarea se evaluará todas las preguntas y no se asignará puntaje a preguntas incompletas.
3. Todo acto de COPIA implica la nota de $0A$. Evita copiar!.

---

1 Preguntas

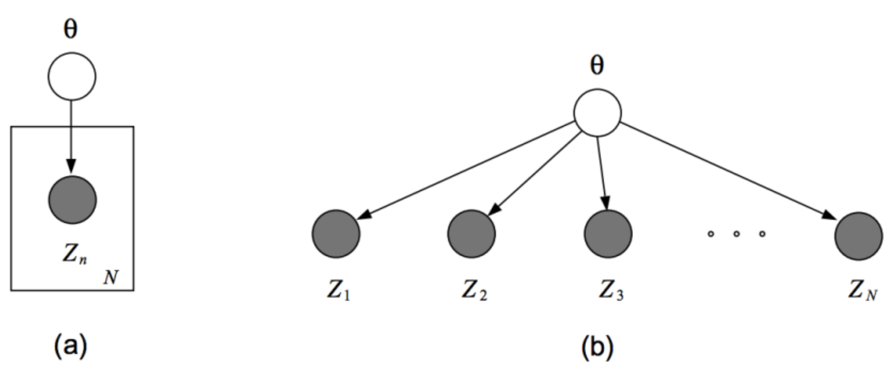
1. Suponga que hay dos HMM que representan dos fonemas: $ph1$ y $ph2$. Cada modelo tiene dos estados y dos observaciones con los siguientes parámetros:

   Modelo 1: $\Pi = [0.5, 0.5], A = [0.5, 0.5|0.5, 0.5], B = [0.8, 0.2|0.2, 0.8]$

   Modelo 2: $\Pi = [0.5, 0.5], A = [0.5, 0.5|0.5, 0.5], B = [0.2, 0.8|0.8, 0.2]$

   Dada la siguiente secuencia de observación: `o1,o1,o2,o2`, ¿cuál es el fonema más probable?.

2. Una representación `lámina` es útil para capturar la replicación en redes Bayesianas. Por ejemplo, la figura 1(a) es una representación equivalente de la figura 1(b). La N en la esquina inferior derecha de la lámina representa el número de réplicas.

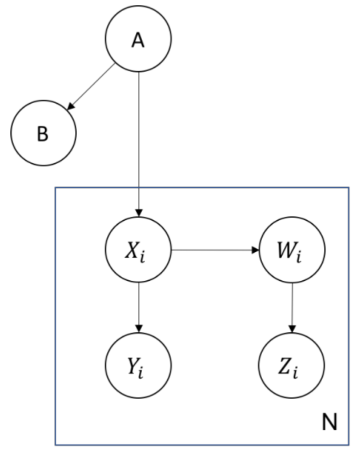


   Ahora considere la red Bayesiana en la figura siguiente. Usamos $X_{1:N}$ como abreviatura de $(X_1, \cdots, X_N)$. Nos gustaría calcular la consulta $P(X_{1:N}|Y_{1:N} = y_{1:N})$. Suponga que todas las variables son binarias.

   (a) ¿Cuál es el número de filas en el factor más grande generado por inferencia por enumeración, para esta consulta?.

   (b) Indica todos los siguientes ordenamientos de eliminación de variables que sean óptimos para calcular la respuesta para la consulta $P(X_{1:N}|Y_{1:N} = y_{1:N})$. (Un orden de eliminación variable es óptimo si los factores más grandes generados son los más pequeños entre todos los posibles ordenamientos de eliminación).

   (c) ¿Cuál de las siguientes variables (W_1, Z_1, A, B) se puede eliminar antes de ejecutar la eliminación de variables, sin afectar el resultado de la inferencia?. Eliminar una variable significa no poner su CPT en nuestro conjunto inicial de factores al iniciar el algoritmo.



- ☐ $Z_1, \cdots, Z_N, W_1, \cdots, W_N, B, A$
- ☐ $W_1, \cdots, W_N, Z_1, \cdots, Z_N, B, A$
- ☐ $A, B, W_1, \cdots, W_N, Z_1, \cdots, Z_N$
- ☐ $A, B, Z_1, \cdots, Z_N, W_1, \cdots, W_N$

3. La mayoría de las técnicas para los procesos de decisión de Markov se centran en calcular $V^*(s)$, la utilidad máxima esperada del estado s. Esta utilidad máxima esperada $V^*(s)$ satisface la siguiente expresión recursiva, conocida como la ecuación de Optimidad de Bellman:

$$V^*(s) = \max_a \sum_{s'} T(s,a,s')[R(s,a,s') + \gamma V^*(s')].$$

En esta pregunta, en lugar de medir la calidad de una política por su utilidad esperada, consideraremos la utilidad del peor de los casos como nuestra medida de calidad. Concretamente, $L^\pi(s)$ es la utilidad mínima que es posible obtener sobre todas las secuencias de acción de estado (potencialmente infinitas) que pueden resultar de la ejecución de la política π a partir del estado s. $L^*(s) = \max_\pi L^\pi(s)$ es la utilidad óptima en el peor de los casos. En palabras, $L^*(s)$ es el mayor límite inferior de la utilidad del estado s.

Sea $C(s,a)$ el conjunto de todos los estados a los que el agente tiene una probabilidad distinta de cero de trasladarse desde el estado s mediante la acción a. Formalmente, $C(s,a) = \{s' | T(s,a,s') > 0\}$. Esta notación puede resultarte útil.

(a) Exprese $L^*(s)$ en una forma recursiva similar a la ecuación de pptimidad de Bellman.

(b) La actualización de Bellman para la iteración de valor es:

$$V_{i+1}(s) \leftarrow \max_a \sum_{s'} T(s,a,s')[R(s,a,s') + \gamma V_i(s')].$$

Define formalmente una actualización similar para calcular $L_{i+1}(s)$ usando L_i.

(c) A partir de este punto, puede asumir que $R(s,a,s') = R(s)$ (las recompensas son una función del estado actual) y que $R(s) \geq 0$ para todos los s. Con estos supuestos, la ecuación de optimidad de Bellman para Q-funciones es:

$$Q^*(s,a) = R(s) + \sum_{s'} T(s,a,s')[\gamma \max_{a'} Q^*(s',a')].$$

Sea $M(s,a)$ el mayor límite inferior de la utilidad del estado s cuando se realiza la acción a (M es para L como Q es para V). (En palabras, si un agente juega de manera óptima después de tomar la acción a desde el estado s, esta es la utilidad que el agente está garantizado a lograr). Define formalmente $M^*(s,a)$ en una forma recursiva similar a cómo se define Q^*.