Lógica epistémica

Diego Tajer

16 de abril de 2017

1. Conocimiento y lógica modal

La lógica epistémica parte de la lógica modal. Un agente conoce un hecho P si P es verdadero en todos los mundos que considera posible (dada su información). Sé P cuando tengo suficiente información como para considerar imposible a -P. Lo representaremos con lógica modal.

2. Lenguaje

Lenguaje

- Conectivos proposicionales
- Sujetos 1, 2, 3... n.
- Proposiciones primiticas: p, p′, q, q′, ...
- Operadores de conocimiento: $K_1, K_2, K_3, ...$

Ejemplo: $K_1K_2p \wedge K_2p$ significa que 1 sabe que 2 sabe que p, y 2 sabe que p.

Un agente n considera posible p exactamente si no sabe $\neg P$. Es decir, si $\neg K_n \neg p$.

'n no sabe si p' = $\neg K_n \neg p \wedge \neg K_n \neg\neg p$ [n considera posible p y considera posible $\neg p$.

3. Modelos

Una estructura de Kripke M para n agentes sobre P es una tupla (S, v, $K_1 ... K_n$), donde:

- S es un conjunto de estados o mundos posibles s
- v es una interpretación que asigna a cada estado s una función de P al conjunto $\{V, F\}$. Es decir, es la función valuación, que determina qué oraciones son verdaderas o falsas en cada estado.
- Cada K_i es una relación binaria en s. Estas relaciones representan las relaciones de posibilidad de acuerdo a un agente en un mundo determinado. $(s, t) \in K_i$ si el agente i considera al mundo t posible, dada su información en s.

- K_i es una relación de equivalencia sobre S. Como tal, es una relación *reflexiva, simétrica y transitiva*:

    - Reflexiva: para todo s, $(s, s) \in K_i$.
    - Simétrica: para todo s y t, si $(s, t) \in K_i$, entonces $(t, s) \in K_i$.
    - Transitiva, para todo s, t y u, si $(s, t) \in K_i$, y $(t, u) \in K_i$, entonces $(s, u) \in K_i$.

  Lo bueno de una relación de equivalencia es que permite particiones (ejemplo: lo que sucede con el predicado diádico ’x es pariente de y’, ’x es aliado de y’ (suponiendo que uno es aliado y pariente de uno mismo, claro), o casos más obvios como ’x tiene la misma edad/nacionalidad/etc. que y’).

Noción de verdad en un mundo en una estructura: (M, s) ⊨ P se lee ’P es verdadero en (M,s)’

(M, s) ⊨ p sii v(s)(p) = V

Las clásulas para los conectivos extensionales son las de siempre. Vamos a lo especial:

$(M, s) \models K_iA$ sii $(M, t) \models A$ para todos los t tales que $(s, t) \in K_i$

Ejercicios sencillos:

- Dado que la relación K_i es reflexiva, ¿qué podemos deducir que sucede en un mundo donde i sabe A?
- Una verdad lógica es una oración verdadera en todo estado. ¿Es posible que un sujeto i no conozca una verdad lógica?
- ¿Puede un sujeto saber algo y sin embargo no saber que lo sabe? [pista: usar transitividad]

Los estados y las relaciones K_i se pueden representar con gráficos.

Dado que todas las K_i son reflexivas y simétricas, podemos graficar sin tomar en cuenta la dirección de las flechas e ignorando los ’loops’. Todo ello se presupone. Los gráficos permiten cierta libertad para representar la información. Por ejemplo, podemos representar un juego de cartas de dos jugadores donde hay tres cartas (A,B y C) y a cada jugador se le da una. Esto es lo que cada jugador sabe en cada situación respecto a las posibles distribución de cartas, representando (A,B) como ’el jugador 1 tiene la carta A, el jugador 2 tiene la B’, etc.

$$
\begin{tikzcd}
(C, B) \arrow[r, dash, "2"] \arrow[d, dash, "1"] & (A, B) \arrow[d, dash, "1"] \\
(C, A) \arrow[d, dash, "2"] & (A, C) \arrow[d, dash, "2"] \\
(B, A) \arrow[r, dash, "1"] & (B, C)
\end{tikzcd}
$$

[Ejercicio: Probar $(M, (A, B) \models K_1(2B \vee 2C)$].

4. Conocimiento común, compartido y distrubuido

Sea G un grupo se sujetos. Agregamos tres expresiones:

- E_G: Todos en G saben que...
- C_G: Es conocimiento común entre los miembros de G que...
- D_G: Es conocimiento distribuido entre los agentes de G que...

Yendo a cosas más serias,

- $(M, s) \models E_G A$ sii $(M, s) \models K_i A$ para todo $i \in G$. Es decir, se da conocimiento compartido de A en G cuando todos los del grupo G saben A.
- $(M, s) \models D_G A$ sii $(M, t) \models A$ para toda t tal que $(s, t) \cap_{i \in G} K_i$. Esto es, se da conocimiento distribuido de A en G cuando, al considerar los estados que *todos* consideran posibles desde s, en todos esos estados A es verdadero. Por eso sólo consideramos la intersección. 'Intuitivamente', es lo que pasaría si la gente se pone a charlar y a realizar aserciones (no es un modelo de sospecha y debate sino de confianza y aumento de información).

Vamos ahora al conocimiento común. $C_G A$ quiere decir que todos en G saben A, pero además saben que todos en G saben, y que saben que saben, etc. Es decir, sea $E_G^0 A$ una abreviación de A, y $E_G^{k+1} A$ una abreviación de $E_G E_G^k A$. Definiremos conocimiento común así:

- $(M, s) \models C_G A$ sii $(M, s) \models E_G^k A$ para k=1,2,3...

Ejercicios:

1. Vuelva a observar el juego de las dos cartas y los tres jugadores. Resuelva lo siguiente:
   - ¿Qué pasa en (A,B) si les digo que C no fue repartida?
   - ¿En (A,B), es verdadero que $E(2B \vee 2C)$?
   - ¿En (A,B), es cierto que $C(2B \vee 2C)$?
   - ¿En (A,B), es cierto que $D(A, B)$?
2. Tamara festeja su cumpleaños. Dado que varios de sus amigos del secundario están peleados entre sí, oculta la lista de destinatarios. En el mail que les envía, avisa que la fecha es el 4 de abril.
   - ¿Hay conocimiento compartido entre sus amigos de que el cumple es el 4 de abril? ¿Y conocimiento común?
   - Supongamos que Tamara se da cuenta del problema, y les manda a cada uno (por separado) un mail con la lista de invitados. ¿Habría ahora conocimiento común?
   - Supongamos que Felicia y Roberto eran novios. Roberto llega al cumpleaños y no la encuentra a Felicia. ¿Puede inferir que Felicia no vino porque sabía que él vendría?
   - Tamara olvida aclarar el lugar del festejo. Su amiga Andrea considera posible que el lugar sea La Cigale o que sea El Dorado. Su amiga Andrea, por otro lado, considera posible tanto El Especial como La Cigale. ¿Hay conocimiento compartido entre Gigi y Andrea del lugar de festejo?

4.1. Alcanzabilidad

Decimos que un estado t es G-alcanzable a partir del estado s en k pasos (1 o más) si hay estados $s_0, s_1, ...s_k$ tales que $s_0 = s$, $s_k = t$ y para todo j con $0 \leq j \leq k-1$, existe un $i \in G$ tal que $(s_j, s_{j+1}) \in K_i$. Decimos que t es G-alcanzable desde s si t es G-alcanzable desde s en 1 o más pasos. Entonces, t es G alcanzable desde s exactamente si hay un camino en el gráfico desde s hacia t cuyos ejes están etiquetados por miembros de G. Cuando G es el conjunto de todos los agentes, decimos directamente que t es alcanzable desde s.

Lemas interesantes:

- $(M, s) \vDash E^k_G A$ sii $(M, t) \vDash A$ para todo t que es G-alcanzable en desde s en k pasos.
- $(M, s) \vDash C_G A$ sii $(M, t) \vDash A$ para todo t que es G-alcanzable en desde s

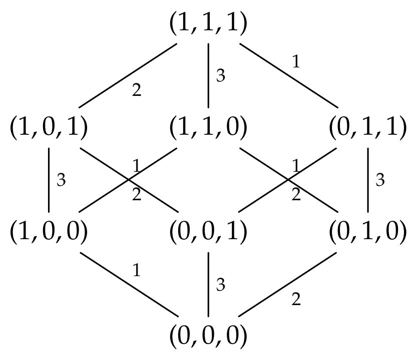
5. Ejemplo: frentes manchadas

Niños: números de 1 a n.

El niño n está manchado = p_n

El niño m sabe que el niño n está manchado = $K_m p_n$

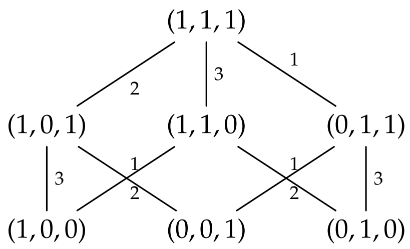
Supongamos que hay tres niños. Llamaremos (0,1,0) a la situación de que el niño 2 está manchado, (0,1,1) a la situación de que los niños 2 y 3 lo están, etc. Podemos graficar el conocimiento así:



Podemos aquí notar algunas cosas. Por ejemplo, que antes de que hable el padre, la oración 'Todos saben que todos saben que hay un niño manchado' ($E^2 p$) es falsa en (1,0,1), ya que la situación (0,0,0) es accesible en dos pasos. Como parece obvio, esa oración solo es verdadera en (1,1,1): allí cada niño sabe que los otros dos niños están manchados, y por ende al verse uno al otro aprenden p.

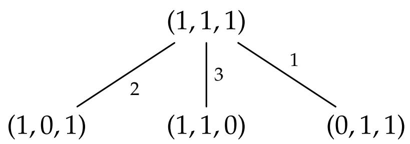
En el momento en que el padre dice que hay un chico manchado, la situación

(0,0,0) deja de ser una posibilidad epistémica. Entonces el gráfico queda así:



Y hay conocimiento común de p, ya que la situación (0,0,0) no es accesible desde ninguna situación.

Cuando el padre pregunta '¿Alguno de ustedes sabe si está manchado?' y los niños dicen 'No', se descartan las situaciones en las que sólo un niño está manchado; pues si así fuera, ese niño hubiera dicho que sabía que él mismo estaba manchado, al ver a los demás limpios. El gráfico queda:



Suponiendo que la situación real fuera (1,0,1), los niños 1 y 3 se darían cuenta que están manchados, ya que uno de sus amiguitos está manchado y el otro no, y han descartado la posibilidad de que sólo uno lo estuviera. En general, luego de k preguntas, es conocimiento común que al menos k+1 niños están manchados.

6. Las propiedades del conocimiento

Llamamos a una oración *válida en M* cuando es verdadera en todo estado s de M.

Llamamos a una oración *satisfacible en M* cuando es verdadera en algún estado s de M.

La llamamos válida *simpliciter* cuando es válida en todas las estructuras.

La llamamos satisfacible *simpliciter* cuando es satisfacible en alguna estructura.

Las estructuras kripkeanas permiten que el conocimiento esté cerrado bajo consecuencia lógica. Es decir, que si conozco A y A implica B, entonces también conozco B.

Algunas otras propiedades que se derivan de nuestra semántica S5:

Factividad o Axioma del conocimiento $\models K_iA \to A$

KK o Introspección positiva $\models K_iA \to K_iK_iA$

Sabiduría o Introspección negativa $\models \neg K_1A \to K_i\neg K_iA$

Regla de la generalización del conocimiento Si $M \models A$, entonces $M \models K_iA$.

Axioma de distribución $\models (K_iA \wedge K_i(A \to B)) \to K_iB$

[Ejercicio: probarlas. Ayuda: factividad proviene de la reflexividad de K, Introspección positiva proviene de su transitividad, Sabiduría proviene de la simetría y la transitividad, Generalización y distribución no dependen de ninguna de las propiedades de K]