

LÓGICA DE CONCIENCIA: BISIMULACIÓN A NIVEL CONCIENCIA, Y CAMBIOS EN CONCIENCIA/CONOCIMIENTO

Fernando R. Velázquez-Quesada
Grupo de Lógica, Lenguaje e Información
Universidad de Sevilla

Universidad de Málaga

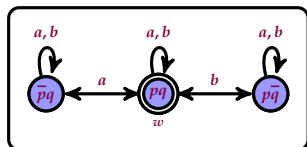
LENGUAJE EPISTÉMICO

DEFINICIÓN (LENGUAJE $\mathcal{L}(\Box)$)

$\varphi ::= \top \mid p \mid \neg\varphi \mid \varphi \vee \psi \mid \Box_i \varphi$ con $p \in \mathbf{P}$ y $i \in \mathbf{A}$.
 $\Box_i \varphi$: "el agente i sabe que φ es verdadera".

DEFINICIÓN (INTERPRETACIÓN SEMÁNTICA)

$(M, w) \models \Box_i \varphi$ ssi para todo $u \in W$, $R_i w u$ implica $(M, u) \models \varphi$



$\rightarrow (M, w) \models \Box_a p$?
 $\rightarrow (M, w) \models \Box_b p$ ✓

MODELO EPISTÉMICO

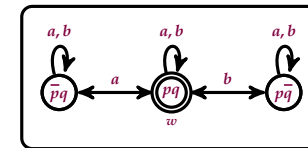
Sea \mathbf{P} un conjunto de proposiciones atómicas; sea \mathbf{A} un conjunto de agentes.

DEFINICIÓN (MODELO EPISTÉMICO)

Un *modelo epistémico* es una tupla $M = \langle W, R_i, V \rangle$ en la cual

- $W \neq \emptyset$ es el conjunto de *mundos posibles*,
- $R_i \subseteq (W \times W)$ es la *relación epistémica* del agente i ,
- $V : \mathbf{P} \rightarrow \wp(W)$ es la *evaluación atómica*.

El par (M, w) , con $w \in W$, es un *estado epistémico*.



¿Y SI LOS AGENTES NO SON IDEALES?

Note como

- la fórmula $\Box_i (\varphi \rightarrow \psi) \rightarrow (\Box_i \varphi \rightarrow \Box_i \psi)$ es válida;
- si φ es válida, entonces $\Box_i \varphi$ es válida;
- todo agente es consciente de toda fórmula del lenguaje;
- ...

En otras palabras, la lógica epistémica representa el conocimiento de agentes *ideales*.

¿Qué sucede con los agentes que no lo son?

En particular, ¿qué sucede con los agentes que no son *conscientes* de todo?

UNA SOLUCIÓN

- $\Box_i \varphi$: "el agente i sabe implícitamente que φ es verdadera", i.e.,

$$\text{Im}_i \varphi := \Box_i \varphi$$

- Para que este conocimiento implícito sea **explícito**, es necesario que el agente esté **consciente** de φ :

$$\text{Ex}_i \varphi := \Box_i \varphi \wedge A_i \varphi$$

- Intuitivamente, un agente que no es consciente de todo **no ve** ciertas posibilidades.¹
- Este trabajo estudia agentes cuya **conciencia/atención** está dada en términos de un **conjunto de proposiciones atómicas**.

¹Esto no es lo mismo que **haber eliminado** algunas posibilidades.

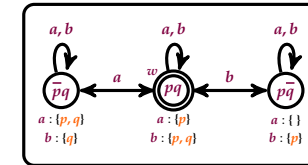
AÑADIENDO CONCIENCIA (1)

DEFINICIÓN (MODELO EPISTÉMICO CON CONCIENCIA)

Un **modelo epistémico con conciencia** es una tupla $M = \langle W, R_i, V, A_i \rangle$ en la cual

- $M = \langle W, R_i, V \rangle$ es un modelo epistémico, y
- $A_i : W \rightarrow \wp(\mathbf{P})$ es la **función de conciencia** del agente i .

El par (M, w) , con $w \in W$, es un **estado epistémico con conciencia**.



AÑADIENDO CONCIENCIA (2)

DEFINICIÓN (LENGUAJE $\mathcal{L}(\Box, A)$)

$$\varphi ::= \top \mid p \mid A_i \varphi \mid \neg \varphi \mid \varphi \vee \psi \mid \Box_i \varphi \quad \text{con } p \in \mathbf{P} \text{ y } i \in \mathbf{A}.$$

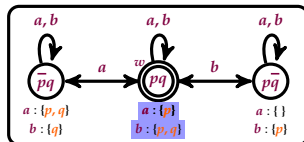
$A_i \varphi$: "el agente i es consciente de φ ".

DEFINICIÓN (INTERPRETACIÓN SEMÁNTICA)

$$(M, w) \models A_i \varphi \quad \text{ssi} \quad v(\varphi) \subseteq A_i(w)$$

donde $v(\varphi)$ son los átomos en φ , i.e.,

$$v(\top) := \emptyset \mid v(p) := \{p\} \mid v(A_i \varphi) := v(\varphi) \mid v(\neg \varphi) := v(\varphi) \mid v(\varphi \wedge \psi) := v(\varphi) \cup v(\psi) \mid v(\Box_i \varphi) := v(\varphi).$$

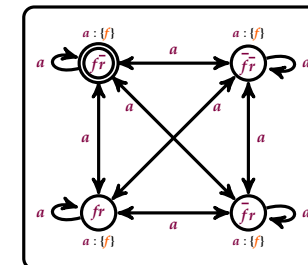


- $(M, w) \models A_a (p \wedge q) \quad ? \times$
- $(M, w) \models A_b (p \wedge q) \quad ? \checkmark$

EJEMPLO

A Alfrid (a) le gusta el futbol. Él es simpatizante de la selección Rusa, y está consciente de que ayer hubo un partido entre Rusia y Austria, pero no sabe quien ganó. A él no le interesan otros deportes, y no está consciente de que ayer también hubo un partido de Rugby entre Rusia y Austria.

f - "Rusia ganó el partido de futbol" r - "Rusia ganó el partido de rugby"



OBSERVE COMO UNA BISIMULACIÓN ... (2)

2.- ... nos permite caracterizar lo que *un agente ideal* puede ver.

PROPOSICIÓN ([10])

Sean (M, w) y (M', w') modelos epistémicos. Si

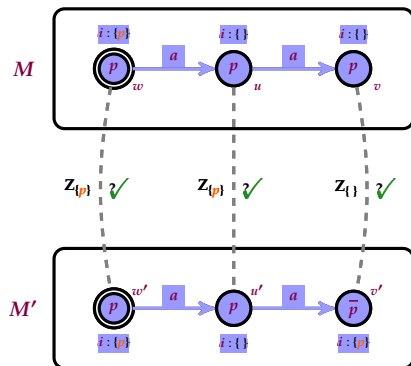
- para cada $u \in R_a[w]$ existe un $u' \in R'_a[w']$ tal que $(M, u) \leftrightarrow_P (M', u')$ y
- para cada $u' \in R'_a[w']$ existe un $u \in R_a[w]$ tal que $(M, u) \leftrightarrow_P (M', u')$,

entonces

$$\{\varphi \mid (M, w) \Vdash \text{Ex}_a \varphi\} = \{\varphi \mid (M', w') \Vdash \text{Ex}_a \varphi\}$$

En particular, si hay un solo agente a y $R_a w w$, entonces \leftrightarrow_P caracteriza lo que a puede ver en (M, w) .

EL CASO EPISTÉMICO CON CONCIENCIA (2)



EL CASO EPISTÉMICO CON CONCIENCIA (1)

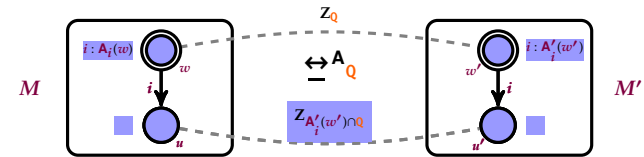
DEFINICIÓN (BISIMULACIÓN A NIVEL CONCIENCIA [8])

Tómese $M = \langle W, R_i, V, A_i \rangle, M' = \langle W', R'_i, V', A'_i \rangle$ y $Q \subseteq P$.

Una Q -bisimulación a nivel conciencia entre M y M' es una relación no vacía $Z_Q \subseteq (W \times W')$ tal que, para todo $(w, w') \in Z_Q$ y $i \in A$,

- (átomos) para todo $p \in Q, w \in V(p)$ ssi $w' \in V'(p)$;
- (conciencia) $A_i(w) \cap Q = A'_i(w') \cap Q$;
- (idaida) si hay un $u \in W$ tal que $R_i w u$, hay un $u' \in W'$ tal que $R'_i w' u'$ y $Z_{A_i(w) \cap Q} u u'$;
- (vuelta) si hay un $u' \in W'$ tal que $R'_i w' u'$, hay un $u \in W$ tal que $R_i w u$ y $Z_{A'_i(w') \cap Q} u u'$.

$(M, w) \leftrightarrow_Q^A (M', w')$: hay una Q -bisimulación a nivel conciencia que contiene a (w, w') .



OBSERVE COMO UNA BISIMULACIÓN CON CONCIENCIA ... (1)

1.- Caracteriza la expresividad de un lenguaje especial.

DEFINICIÓN (LENGUAJE $\mathcal{L}(\text{Ex})$)

$$\varphi ::= \top \mid p \mid \neg\varphi \mid \varphi \vee \psi \mid \text{Ex}_i \varphi \quad \text{con } p \in P \text{ y } i \in A.$$

PROPOSICIÓN ([10])

Sean (M, w) y (M', w') modelos epistémicos con conciencia.

- $(M, w) \leftrightarrow_Q^A (M', w') \implies (M, w) \equiv_Q^{\mathcal{L}(\text{Ex})} (M', w')$
- Si M y M' tienen imagen finita, entonces $(M, w) \leftrightarrow_Q^A (M', w') \iff (M, w) \equiv_Q^{\mathcal{L}(\text{Ex})} (M', w')$

Es decir, \leftrightarrow_Q^A caracteriza lo que el lenguaje $\mathcal{L}(\text{Ex})_Q$ puede ver.

OBSERVE COMO UNA BISIMULACIÓN CON CONCIENCIA . . . (2)

2.- Nos permite caracterizar lo que **un agente** puede ver.

PROPOSICIÓN ([10])

Sean (M, w) y (M', w') modelos epistémicos con conciencia. Si

- $A_a(w) = A'_a(w')$,
- para cada $u \in R_a[w]$ existe un $u' \in R'_a[w']$ tal que $(M, u) \stackrel{A}{\sim}_{A_a(w)} (M', u')$ y
- para cada $u' \in R'_a[w']$ existe un $u \in R_a[w]$ tal que $(M, u) \stackrel{A}{\sim}_{A'_a(w')} (M', u')$,

entonces

$$\{\varphi \mid (M, w) \models \text{Ex}_a \varphi\} = \{\varphi \mid (M', w') \models \text{Ex}_a \varphi\}$$

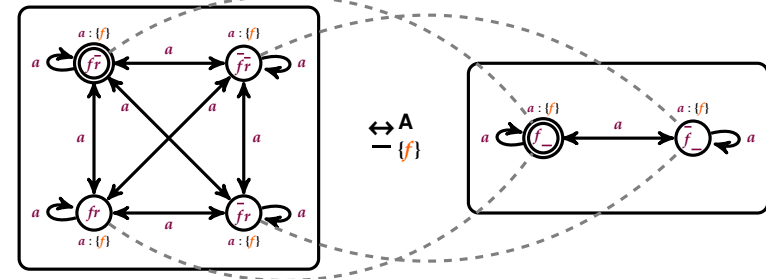
En particular, si hay un solo agente a y $R_a w w$, entonces $\stackrel{A}{\sim}_{A_a(w)}$ caracteriza **lo que a puede ver en (M, w)** .

ACCIONES EPISTÉMICAS

- Llega nueva información (anuncios, observaciones).
- Cambios en conciencia (revelar u ocultar posibilidades).

EJEMPLO

A Alfrid (a) le gusta el fútbol. Él es simpatizante de la selección Rusa, y está consciente de que ayer hubo un partido entre Rusia y Austria, pero no sabe quien ganó. A él no le interesan otros deportes, y no está consciente de que ayer también hubo un partido de Rugby entre Rusia y Austria.



¿CÓMO REPRESENTAR ESTAS ACCIONES?

- Versiones públicas:
 - anuncios públicos: eliminar mundos posibles del modelo [3, 2].
 - cambios públicos de conciencia: agregar/quitar átomos al/del conjunto de conciencia de todos los agentes [8].
- Versiones privadas:
 - anuncios privados: eliminar mundos de entre las posibilidades de algunos agentes, manteniendo la estructura anterior para el resto de ellos [1].
 - cambios privados de conciencia: agregar/quitar átomos al/del conjunto de conciencia de algunos agentes, manteniendo la estructura anterior para el resto de ellos [7].

MODELOS DE ACCIÓN CON CONCIENCIA

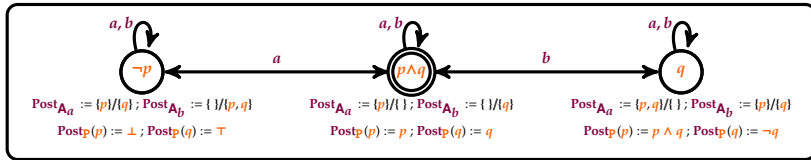
La incertidumbre acerca de la acción se representa como la incertidumbre acerca de la situación [1].

DEFINICIÓN (MODELOS DE ACCIÓN CON CONCIENCIA [6, 9])

Un *modelo de acción con conciencia* es una tupla $\mathbf{M} = \langle D, T_i, \text{Pre}, \text{Post}_{A_i}, \text{Post}_p \rangle$ en la cual

- $D \neq \emptyset$ es un conjunto de *eventos posibles*,
- $T_i \subseteq (W \times W)$ es la *relación epistémica* del agente i ,
- $\text{Pre} : D \rightarrow \mathcal{L}$ es la *función de precondition* (con \mathcal{L} algún lenguaje),
- $\text{Post}_{A_i} : \{+, -\} \rightarrow D \rightarrow \wp(\mathcal{P})$ es la *función de cambio de conciencia*,
- $\text{Post}_p : D \rightarrow \mathcal{P} \rightarrow \mathcal{L}$ es la *función de postcondición*.

El par (\mathbf{M}, d) , con $d \in D$, es una *acción con conciencia*.



EL LENGUAJE

Agregamos la modalidad $\langle \mathbf{M}, d \rangle$.

DEFINICIÓN (INTERPRETACIÓN SEMÁNTICA)

$$(M, w) \Vdash \langle \mathbf{M}, d \rangle \varphi \quad \text{ssi} \quad (M, w) \Vdash \text{Pre}(d) \text{ y } (M \otimes \mathbf{M}, (w, d)) \Vdash \varphi$$

ACTUALIZACIÓN

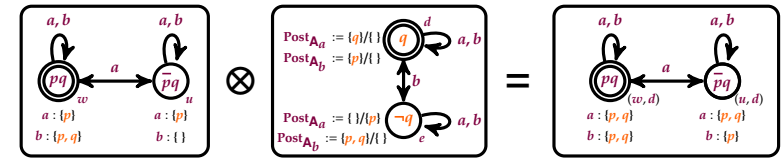
¿Cómo afecta un modelo de acción con conciencia a un modelo epistémico con conciencia?

DEFINICIÓN (ACTUALIZACIÓN)

Sea $M = \langle W, R_i, V, A_i \rangle$ un MEC y $\mathbf{M} = \langle D, T_i, \text{Pre}, \text{Post}_{A_i}, \text{Post}_p \rangle$ un MAC.

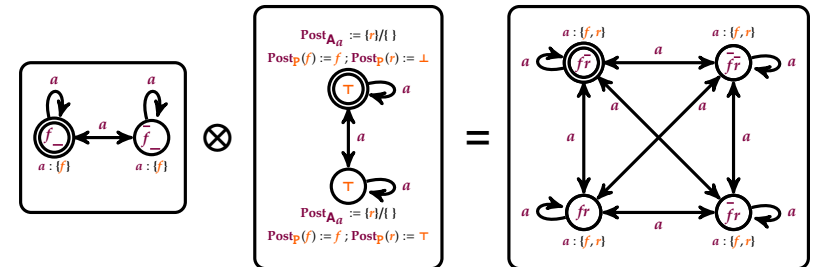
El modelo epistémico con conciencia $M \otimes \mathbf{M} = \langle W^\otimes, R_i^\otimes, V^\otimes, A_i^\otimes \rangle$ se define como

- $W^\otimes := \{(w, d) \in (W \times D) \mid (M, w) \Vdash \text{Pre}(d)\}$;
- $R_i^\otimes := \{((w, d), (u, e)) \in (W^\otimes \times W^\otimes) \mid R_i w u \text{ y } T_i d e\}$;
- $V^\otimes(p) := \{(w, d) \in W^\otimes \mid (M, w) \Vdash \text{Post}_p(d)(p)\}$.
- $A_i^\otimes(w, d) := (A_i(w) \cup \text{Post}_{A_i}^+(d)) \setminus \text{Post}_{A_i}^-(d)$;



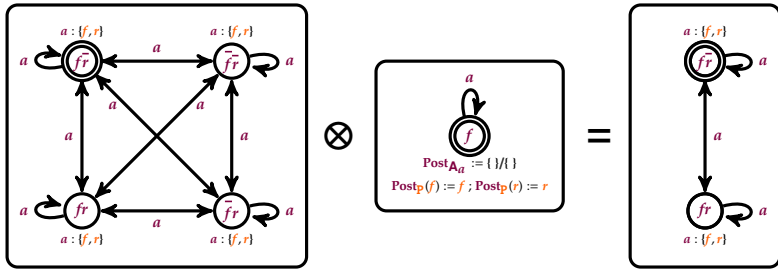
EJEMPLO (1)

Mientras busca el marcador final del partido de fútbol en línea, Alfrid ve una página web con el encabezado “El equipo Ruso de rugby tuvo un partido difícil ayer”, enterándose así que hubo un partido de rugby (pero sin enterarse que Rusia lo perdió).



EJEMPLO (2)

Alfrid sigue buscando el marcador del partido de futbol, y finalmente lo encuentra: Rusia 2 - Austria 1.



- En lógica de conciencia, $Aw \varphi := A \varphi$, pero podemos definir
 - $Aw \varphi := \Box A \varphi$, o
 - $Aw \varphi := [*]A \varphi$, entre otros.
- Comparar estos modelos con otros utilizados para representar conceptos similares, tales como
 - el concepto de *tema en discusión* (issue: [4]),
 - el concepto de *ideas a discutir* (agenda: [5]).

¿POR QUÉ MODELOS DE ACCIÓN CON CONCIENCIA?

TEOREMA ([10])

Sean (M, w) y (M', w') estados epistémicos con conciencia (finitos). Hay una acción con conciencia (M, d) tal que $(M, w) \otimes (M, d)$ es bisimilar a (M', w') .

- A. Baltag, L. S. Moss, and S. Solecki. The logic of public announcements, common knowledge and private suspicions. Technical Report SEN-R9922, CWI, Amsterdam, 1999.
- J. Gerbrandy and W. Groeneveld. Reasoning about information change. *Journal of Logic, Language, and Information*, 6(2):147–196, 1997.
- J. A. Plaza. Logics of public communications. In M. L. Emrich, M. S. Pfeifer, M. Hadzikadic, and Z. W. Ras, editors, *Proceedings of the 4th International Symposium on Methodologies for Intelligent Systems*, pages 201–216, Tennessee, USA, 1989. Oak Ridge National Laboratory, ORNL/DSRD-24.
- J. van Benthem and Ștefan Minică. Toward a dynamic logic of questions. *Journal of Philosophical Logic*, 41(4):633–669, 2012.
- J. van Benthem, P. Girard, and O. Roy. Everything else being equal: A modal logic for ceteris paribus preferences. *Journal of Philosophical Logic*, 38(1):83–125, 2009.
- J. van Benthem and F. R. Velázquez-Quesada. The dynamics of awareness. *Synthese (Knowledge, Rationality and Action)*, 177(Supplement 1):5–27, Dec. 2010.
- H. van Ditmarsch and T. French. Becoming aware of propositional variables. In M. Banerjee and A. Seth, editors, *ICLA*, volume 6521 of *Lecture Notes in Computer Science*, pages 204–218. Springer, 2011.
- H. P. van Ditmarsch and T. French. Awareness and forgetting of facts and agents. In *Web Intelligence/LAT Workshops*, pages 478–483. IEEE, 2009.
- H. P. van Ditmarsch, T. French, and F. R. Velázquez-Quesada. Action models for knowledge and awareness. In W. van der Hoek, L. Padgham, V. Conitzer, and M. Winikoff, editors, *AAMAS*, pages 1091–1098. IFAAMAS, 2012.
- H. P. van Ditmarsch, T. French, F. R. Velázquez-Quesada, and Y. Wang. Knowledge, awareness, and bisimulation. In B. C. Schipper, editor, *TARK*, pages 61–70, 2013.