Una lógica epistémica para el razonamiento estratégico

Andrés R. Saravia

Facultad de Matemática, Astronomía, Física y Computación, Universidad Nacional de Córdoba

14 de octubre de 2021



Doctorando en Ciencias de la Computación (desde 2020)



- Doctorando en Ciencias de la Computación (desde 2020)
- Director: Raul Fervari



- Doctorando en Ciencias de la Computación (desde 2020)
- Director: Raul Fervari
- Trabajamos en lógicas Knowing How o "saber cómo" desde una perspectiva epistémica:



Knowing That

- Doctorando en Ciencias de la Computación (desde 2020)
- Director: Raul Fervari
- Trabajamos en lógicas Knowing How o "saber cómo" desde una perspectiva epistémica:
 - lenguajes epistémicos centrados en describir el conocimiento de los agentes sobre sus habilidades.

Knowing That

- Doctorando en Ciencias de la Computación (desde 2020)
- Director: Raul Fervari
- Trabajamos en lógicas Knowing How o "saber cómo" desde una perspectiva epistémica:
 - lenguajes epistémicos centrados en describir el conocimiento de los agentes sobre sus habilidades.
- Resultado: Paper aceptado en TARK 2021

Lógica Knowing That: Lógica Modal Epistémica



- Lógica Knowing That: Lógica Modal Epistémica
- Lógica Knowing How: Propuesta sobre modelos LTS



- Lógica Knowing That: Lógica Modal Epistémica
- Lógica Knowing How: Propuesta sobre modelos LTS
- Nuestra propuesta: Knowing how sobre modelos LTS^U



- Lógica Knowing That: Lógica Modal Epistémica
- Lógica Knowing How: Propuesta sobre modelos LTS
- Nuestra propuesta: Knowing how sobre modelos LTS^U
- Conclusiones y trabajos futuros

Knowing That

Consideremos dos agentes: Ana (a) y Bruno (b); Ana está en el parque y Bruno está en su casa con las ventanas cerradas.



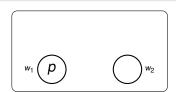
Knowing That

Consideremos dos agentes: Ana (a) y Bruno (b); Ana está en el parque y Bruno está en su casa con las ventanas cerradas. Tenemos que Ana sabe que está soleado. Pero Bruno no.

Knowing That

Consideremos dos agentes: Ana (a) y Bruno (b); Ana está en el parque y Bruno está en su casa con las ventanas cerradas. Tenemos que Ana sabe que está soleado. Pero Bruno no.

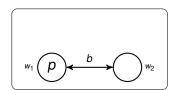
p: está soleado



Knowing That

Consideremos dos agentes: Ana (a) y Bruno (b); Ana está en el parque y Bruno está en su casa con las ventanas cerradas. Tenemos que Ana sabe que está soleado. Pero Bruno no.

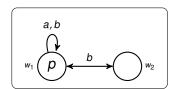
p: está soleado



Knowing That

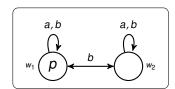
Consideremos dos agentes: Ana (a) y Bruno (b); Ana está en el parque y Bruno está en su casa con las ventanas cerradas. Tenemos que Ana sabe que está soleado. Pero Bruno no.

p: está soleado



Consideremos dos agentes: Ana (a) y Bruno (b); Ana está en el parque y Bruno está en su casa con las ventanas cerradas. Tenemos que Ana sabe que está soleado. Pero Bruno no.

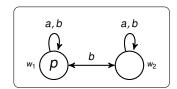
p: está soleado



Knowina How: LTS^Us

Consideremos dos agentes: Ana (a) y Bruno (b); Ana está en el parque y Bruno está en su casa con las ventanas cerradas. Tenemos que Ana sabe que está soleado. Pero Bruno no.

p: está soleado



El conocimiento de Ana y Bruno sobre p se puede escribir en esta lógica como $K_a p$ y $\neg K_b p$.



Dado un agente i y su relación de indistinguibilidad entre mundos \sim_i . En la Lógica Modal Epistémica, el conocimiento del agente cumple ciertas propiedades.

• Verdad y conocimiento: $K_i \varphi \to \varphi$.

Dado un agente i y su relación de indistinguibilidad entre mundos \sim_i . En la Lógica Modal Epistémica, el conocimiento del agente cumple ciertas propiedades.

Knowina How: LTS^Us

- Verdad y conocimiento: $K_i \varphi \to \varphi$.
- Introspección positiva: $K_i \varphi \to K_i K_i \varphi$.

- Verdad y conocimiento: $K_i \varphi \to \varphi$.
- Introspección positiva: $K_i \varphi \to K_i K_i \varphi$.
- Introspección negativa: $\neg K_i \varphi \rightarrow K_i \neg K_i \varphi$.

- Verdad y conocimiento: $K_i \varphi \to \varphi$.
 - ~; reflexiva.
- Introspección positiva: $K_i \varphi \to K_i K_i \varphi$.
- Introspección negativa: $\neg K_i \varphi \rightarrow K_i \neg K_i \varphi$.

Knowing That

- Verdad y conocimiento: $K_i \varphi \to \varphi$.
 - ~; reflexiva.
- Introspección positiva: $K_i \varphi \to K_i K_i \varphi$.
 - ~i transitiva.
- Introspección negativa: $\neg K_i \varphi \rightarrow K_i \neg K_i \varphi$.

- Verdad y conocimiento: $K_i \varphi \to \varphi$.
 - ~; reflexiva.
- Introspección positiva: $K_i \varphi \to K_i K_i \varphi$.
 - ~i transitiva.
- Introspección negativa: $\neg K_i \varphi \rightarrow K_i \neg K_i \varphi$.
 - ~; euclídea.

Dado un agente i y su relación de indistinguibilidad entre mundos \sim_i . En la Lógica Modal Epistémica, el conocimiento del agente cumple ciertas propiedades.

Knowina How: LTSUs

- Verdad y conocimiento: $K_i \varphi \to \varphi$.
 - ~; reflexiva.
- Introspección positiva: $K_i \varphi \to K_i K_i \varphi$.
 - ~i transitiva.
- Introspección negativa: $\neg K_i \varphi \rightarrow K_i \neg K_i \varphi$.
 - ∘ ~; euclídea.
 - Las anteriores implican ~_i simétrica.

Dado un agente i y su relación de indistinguibilidad entre mundos \sim_i . En la Lógica Modal Epistémica, el conocimiento del agente cumple ciertas propiedades.

Knowing How: LTSUs

- Verdad y conocimiento: $K_i \varphi \to \varphi$.
 - ∼_i reflexiva.
- Introspección positiva: $K_i \varphi \to K_i K_i \varphi$.
 - ~i transitiva.
- Introspección negativa: $\neg K_i \varphi \rightarrow K_i \neg K_i \varphi$.
 - ~; euclídea.
 - Las anteriores implican ~_i simétrica.

Semánticamente, el conocimiento de un agente es una relación de equivalencia.

Otros patrones del conocimiento:



Otros patrones del conocimiento: knowing why



Otros patrones del conocimiento: knowing why, knowing whether



 Otros patrones del conocimiento: knowing why, knowing whether, knowing who



• Otros patrones del conocimiento: knowing why, knowing whether, knowing who y knowing how.



- Otros patrones del conocimiento: knowing why, knowing whether, knowing who y knowing how.
- Se centra en las habilidades de los agentes.



- Otros patrones del conocimiento: knowing why, knowing whether, knowing who y knowing how.
- Se centra en las habilidades de los agentes.
- Wang propuso un marco para las lógicas knowing how.

Y. Wang (2015): A Logic of Knowing How. (LORI 2015).



Knowing How: LTSUs

$$S = \langle W, \{R_a\}_{a \in Act}, V \rangle$$
 donde

Knowing How: LTSUs

 $S = \langle W, \{R_a\}_{a \in Act}, V \rangle$ donde

• W es un conjunto de estados,

Knowing How: LTSUs

 $S = \langle W, \{R_a\}_{a \in Act}, V \rangle$ donde

- W es un conjunto de estados,
- $R_a \subseteq W \times W$ por cada $a \in Act$,

Sistemas de transiciones etiquetadas (LTSs)

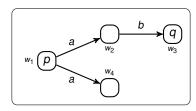
 $S = \langle W, \{R_a\}_{a \in Act}, V \rangle$ donde

- W es un conjunto de estados,
- $R_a \subseteq W \times W$ por cada $a \in Act$,
- V : Prop $\rightarrow 2^{W}$.

Sistemas de transiciones etiquetadas (LTSs)

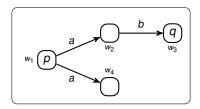
 $S = \langle W, \{R_a\}_{a \in Act}, V \rangle$ donde

- W es un conjunto de estados,
- $R_a \subseteq W \times W$ por cada $a \in Act$,
- V : Prop $\rightarrow 2^{W}$.



 $S = \langle W, \{R_a\}_{a \in Act}, V \rangle$ donde

- W es un conjunto de estados,
- $R_a \subseteq W \times W$ por cada $a \in Act$,
- V : Prop $\rightarrow 2^{W}$.

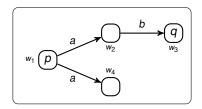


Representa las acciones que puede ejecutar el agente.

Sistemas de transiciones etiquetadas (LTSs)

 $S = \langle W, \{R_a\}_{a \in Act}, V \rangle$ donde

- W es un conjunto de estados,
- $R_a \subseteq W \times W$ por cada $a \in Act$,
- V : Prop $\rightarrow 2^{W}$.



Knowina How: LTS^Us

Representa las acciones que puede ejecutar el agente. Una transición a de w₁ a w₂ se interpreta como "luego de ejecutar la acción a en el estado w_1 el agente llega al estado w_2 ".

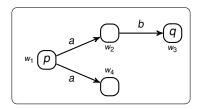


Sistemas de transiciones etiquetadas (LTSs)

 $S = \langle W, \{R_a\}_{a \in Act}, V \rangle$ donde

- W es un conjunto de estados,
- $R_a \subseteq W \times W$ por cada $a \in Act$,
- V : Prop $\rightarrow 2^{W}$.

Knowing That



Representa las acciones que puede ejecutar el agente. Una transición a de w₁ a w₂ se interpreta como "luego de ejecutar la acción a en el estado w_1 el agente llega al estado w_2 ". Dado un conjunto Act, un plan σ es un elemento de Act* (a, ab y ϵ).

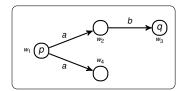
Un plan debe ser a prueba de fallos:

Ejecutabilidad fuerte

Un plan debe ser a prueba de fallos: Cada ejecución parcial debe ser completada.

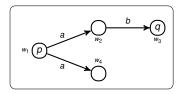


Un plan debe ser a prueba de fallos: Cada ejecución parcial debe ser completada. Ejemplo: el plan ab.



Ejecutabilidad fuerte

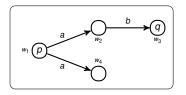
Un plan debe ser a prueba de fallos: Cada ejecución parcial debe ser completada. Ejemplo: el plan *ab*.



ab no es fuertemente ejecutable en *w*₁

Ejecutabilidad fuerte

Un plan debe ser a prueba de fallos: Cada ejecución parcial debe ser completada. Ejemplo: el plan *ab*.



ab no es fuertemente ejecutable en w_1

Definición (Ejecutabilidad fuerte de un plan)

Un plan σ es fuertemente ejecutable (FE) en un $u \in W$ sii para toda ejecución parcial de σ desde u se completa.



L_{Kh} sobre LTS

Definición (Sintaxis L_{Kh})

$$\varphi ::= p \mid \neg \varphi \mid \varphi \lor \varphi \mid \mathsf{Kh}(\varphi, \varphi)$$

 $\mathsf{Kh}(\psi,\varphi)$: "cuando ψ se cumple, el agente sabe cómo hacer φ verdadera".

L_{Kh} sobre LTS

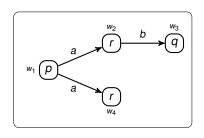
Definición (Sintaxis L_{Kh})

$$\varphi ::= p \mid \neg \varphi \mid \varphi \lor \varphi \mid \mathsf{Kh}(\varphi, \varphi)$$

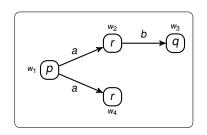
Kh(ψ , φ): "cuando ψ se cumple, el agente sabe cómo hacer φ verdadera".

Definición (L_{Kh} en LTS)

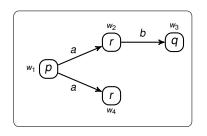
- S, $w \models \mathsf{Kh}(\psi, \varphi)$ sii_{def} existe un plan $\sigma \in \mathsf{Act}^*$ tal que
 - (1) σ es FE en todos los ψ -mundos y
 - (2) desde los ψ -mundos, σ siempre termina en φ -mundos.



$$S$$
, $w_1 \models Kh(p, r)$

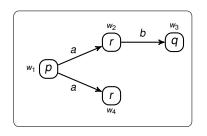


 $S, w_1 \models \mathsf{Kh}(p, r)$ el plan a es FE en w_1 (p-mundo), y lleva de p-mundos a r-mundos.

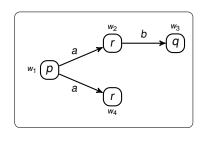


 $S, w_1 \models \mathsf{Kh}(p, r)$ el plan a es FE en w_1 (p-mundo), y lleva de p-mundos a r-mundos.

$$S, w_1 \not\models \mathsf{Kh}(p, q)$$



- $S, w_1 \models \mathsf{Kh}(p, r)$ el plan a es FE en w_1 (p-mundo), y lleva de p-mundos a r-mundos.
- $S, w_1 \not\models \mathsf{Kh}(p,q)$ - ϵ y a: son FE en w_1 (p-mundo), pero no llevan a q;



 $S, w_1 \models \mathsf{Kh}(p, r)$ el plan a es FE en w_1 (p-mundo), y lleva de p-mundos a r-mundos.

Knowing How: LTSUs

 $S, w_1 \not\models \mathsf{Kh}(p,q)$ - ϵ y a: son FE en w_1 (p-mundo), pero no llevan a q; - ab no es FE en w₁.

Una lógica 'no tan' epistémica

• Saber qué:

- información óntica: hechos o proposiciones verdaderas en un mundo posible;
- información epistémica: relación de incertidumbre o indistinguibilidad, percepción del agente.



Una lógica 'no tan' epistémica

Saber qué:

- información óntica: hechos o proposiciones verdaderas en un mundo posible;
- información epistémica: relación de incertidumbre o indistinguibilidad, percepción del agente.

Saber cómo:

- el agente tiene disponible todos los planes para elegir un testigo.
- o no hay distinción entre óntico y epistémico.



Nos gustaría modelar la incertidumbre

• En estos modelos el agente tiene disponible todos los planes para elegir un testigo.



Nos gustaría modelar la incertidumbre

• En estos modelos el agente tiene disponible todos los planes para elegir un testigo.

¿Y si no tuviera el conocimiento de que existen ciertos planes?

 En estos modelos el agente tiene disponible todos los planes para elegir un testigo.

Knowina How: LTS^Us

¿Y si no tuviera el conocimiento de que existen ciertos planes?

Ejemplo: el plan vacío ("hacer nada").

Nos gustaría modelar la incertidumbre

• En estos modelos el agente tiene disponible todos los planes para elegir un testigo.

¿Y si no tuviera el conocimiento de que existen ciertos planes?

Ejemplo: el plan vacío ("hacer nada").

Para el agente cada plan es distinto de cualquier otro.

• En estos modelos el agente tiene disponible todos los planes para elegir un testigo.

Knowina How: LTS^Us

- ¿Y si no tuviera el conocimiento de que existen ciertos planes?
 - Ejemplo: el plan vacío ("hacer nada").
- Para el agente cada plan es distinto de cualquier otro.
 - ¿Y si no fuera capaz de distinguir ciertos planes de otros?

- En estos modelos el agente tiene disponible todos los planes para elegir un testigo.
 - ¿Y si no tuviera el conocimiento de que existen ciertos planes?
 - Ejemplo: el plan vacío ("hacer nada").
- Para el agente cada plan es distinto de cualquier otro.
 - ¿Y si no fuera capaz de distinguir ciertos planes de otros?

Hay diferentes razones para no saber cómo. El agente no puede distinguir entre acciones básicas, el orden, etc.

Nos gustaría modelar la incertidumbre

 En estos modelos el agente tiene disponible todos los planes para elegir un testigo.

¿Y si no tuviera el conocimiento de que existen ciertos planes?

Ejemplo: el plan vacío ("hacer nada").

Para el agente cada plan es distinto de cualquier otro.

¿Y si no fuera capaz de distinguir ciertos planes de otros?

Hay diferentes razones para no saber cómo. El agente no puede distinguir entre acciones básicas, el orden, etc.

C. Areces, R. Fervari, A. Saravia, F. Velázquez-Quesada. *Uncertainty-Based Semantics for Multi-Agent Knowing How Logics.* (TARK 2021).



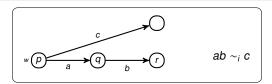
LTS basado en Incertidumbre (LTS^U)

Definición (LTS basado en Incertidumbre)

Dado un conjunto finito no vacío Agt, un LTSU es una tupla $\mathcal{M} = \langle W, \{R_a\}_{a \in Act}, \{\sim_i\}_{i \in Act}, V \rangle$ donde

- $\langle W, \{R_a\}_{a \in Act}, V \rangle$ es un LTS,
- ∼i es una relación de equivalencia (indistinguibilidad) sobre un conjunto no vacío de planes.

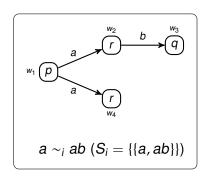
Knowina How: LTS^Us



Definición (L_{Kh}, en LTS^U)

Sea S_i el conjunto de clases de equivalencia (sobre planes) por \sim_i . \mathcal{M} , $w \models \mathsf{Kh}_i(\psi, \varphi)$ sii existe un conjunto de planes $\pi \in S_i$ tal que

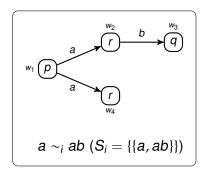
- (1) Todos los planes en π son FE en todos los ψ -mundos y
- (2) desde los ψ -mundos, cada plan de π siempre termina en φ -mundos.



$$\mathcal{M}, \mathbf{w}_1 \models \neg \mathsf{Kh}_i(p, r)$$

Knowing How: LTSUs

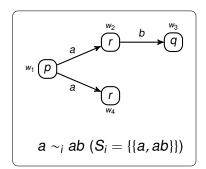
0000



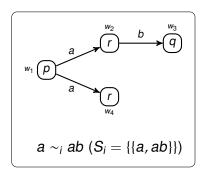
$$\mathcal{M}$$
, $w_1 \models \neg \mathsf{Kh}_i(p,r)$
el plan a es FE en w_1 (p -mundo),
y lleva de mundos p a r -mundos.

Knowing How: LTSUs

0000



 $\mathcal{M}, w_1 \models \neg \mathsf{Kh}_i(p, r)$ el plan a es FE en w_1 (p-mundo), y lleva de mundos p a r-mundos. el plan ab no es FE en w₁ (p-mundo),



 $\mathcal{M}, w_1 \models \neg \mathsf{Kh}_i(p, r)$ el plan a es FE en w_1 (p-mundo), y lleva de mundos p a r-mundos. el plan ab no es FE en w₁ (p-mundo), y el conjunto {a, ab} no es FE en w₁.

Knowina How: LTS^Us

0000

Complejidad

Proposición

Si φ es una fórmula de L_{Kh}, satisfacible, entonces existen un modelo \mathcal{M}' polinomial en el tamaño de φ y un estado w tal que \mathcal{M}' , w $\models \varphi$.

Proposición

Si φ es una fórmula de L_{Kh_i} satisfacible, entonces existen un modelo \mathcal{M}' polinomial en el tamaño de φ y un estado w tal que \mathcal{M}' , w $\models \varphi$.

Knowina How: LTS^Us

0000

Teorema

- Model checking para LKhi está en P.
- El problema de satisfacibilidad para L_{Kh}, es NP-completo.



Resumen

• Lógicas Knowing How Multiagente basadas en Incertidumbre:



Resumen

- Lógicas Knowing How Multiagente basadas en Incertidumbre:
 - Indistinguibilidad entre planes, para múltiples agentes.



Knowing How: LTSUs

- o Indistinguibilidad entre planes, para múltiples agentes.
- o Describe otras razones de no "saber cómo".

Knowing How: LTSUs

- Indistinguibilidad entre planes, para múltiples agentes.
- o Describe otras razones de no "saber cómo".
- Model checking está en P.

Resumen

Lógicas Knowing How Multiagente basadas en Incertidumbre:

Knowina How: LTS^Us

- Indistinguibilidad entre planes, para múltiples agentes.
- Describe otras razones de no "saber cómo".
- Model checking está en P.
- El problema de satisfacibilidad es NP-completo.



Knowina How: LTS^Us

- o Indistinguibilidad entre planes, para múltiples agentes.
- o Describe otras razones de no "saber cómo".
- Model checking está en P.
- El problema de satisfacibilidad es NP-completo.

Además:

Resumen

Lógicas Knowing How Multiagente basadas en Incertidumbre:

Knowina How: LTS^Us

- Indistinguibilidad entre planes, para múltiples agentes.
- Describe otras razones de no "saber cómo".
- Model checking está en P.
- El problema de satisfacibilidad es NP-completo.

Además:

Sistema axiomático fuertemente completo.



Knowina How: LTS^Us

- Indistinguibilidad entre planes, para múltiples agentes.
- o Describe otras razones de no "saber cómo".
- Model checking está en P.
- El problema de satisfacibilidad es NP-completo.

Además:

- Sistema axiomático fuertemente completo.
- Se puede recapturar la lógica basada en LTS en una clase de modelos LTS^U.



- Lógicas Knowing How Multiagente basadas en Incertidumbre:
 - Indistinguibilidad entre planes, para múltiples agentes.

Knowina How: LTS^Us

- o Describe otras razones de no "saber cómo".
- Model checking está en P.
- El problema de satisfacibilidad es NP-completo.

Además:

- Sistema axiomático fuertemente completo.
- Se puede recapturar la lógica basada en LTS en una clase de modelos LTS^U.
- o Una lógica débil pero más general.



Combinar las modalidades knowing how + knowing that,



- Combinar las modalidades knowing how + knowing that,
- Investigar modalidades dinámicas para learning/forgetting how.



- Combinar las modalidades knowing how + knowing that,
- Investigar modalidades dinámicas para learning/forgetting how.
- Nociones de conocimiento colectivo, grupos de agentes.



- Combinar las modalidades knowing how + knowing that,
- Investigar modalidades dinámicas para learning/forgetting how.
- Nociones de conocimiento colectivo, grupos de agentes.
- Aprovechando la flexibilidad de nuestro framework:



Knowing That

- Combinar las modalidades knowing how + knowing that,
- Investigar modalidades dinámicas para learning/forgetting how.
- Nociones de conocimiento colectivo, grupos de agentes.
- Aprovechando la flexibilidad de nuestro framework:
 - o Otras clases de modelos.

Knowing That

- Combinar las modalidades knowing how + knowing that,
- Investigar modalidades dinámicas para learning/forgetting how.
- Nociones de conocimiento colectivo, grupos de agentes.
- Aprovechando la flexibilidad de nuestro framework:
 - Otras clases de modelos.
 - Diferentes condiciones de ejecutabilidad.

Knowing That

- Combinar las modalidades knowing how + knowing that,
- Investigar modalidades dinámicas para learning/forgetting how.
- Nociones de conocimiento colectivo, grupos de agentes.
- Aprovechando la flexibilidad de nuestro framework:
 - o Otras clases de modelos.
 - o Diferentes condiciones de ejecutabilidad.
 - Otros axiomas.