

Lógica Epistémica

Andrés Laurito

Primer cuatrimestre 2016

andy.laurito@hotmail.com

June 29, 2016

Lo que vamos a ver

1 Introducción al problema

- Los papers
- Lógica epistémica

2 El problema - Parte 1

- Definiendo la semántica
- Modelando la noción de poder de razonamiento
- Estudiando la complejidad computacional

Qué papers elegí para presentar?

- ① On the Complexity of Epistemic Reasoning
- ② Which Semantics for Neighbourhood Semantics?

Idea de hacia donde vamos

Tratan el problema de la lógica epistémica con dos enfoques distintos. En uno nos enfocamos en la complejidad computacional de SAT para la lógica modal, mientras que en el otro se trata de dar un nuevo enfoque sobre el mismo problema.

Qué es la lógica epistémica?

Wikipedia

La lógica epistémica es un campo de la lógica modal que se ocupa del razonamiento sobre el conocimiento. Está tiene aplicaciones en numerosos campos, tales como filosofía, ciencia computacional teórica, inteligencia artificial, economía y lingüística.

Stanford Encyclopedia of Philosophy

Epistemic logic is the logic of knowledge and belief. It provides insight into the properties of individual knowers, has provided a means to model complicated scenarios involving groups of knowers and has improved our understanding of the dynamics of inquiry.

Qué es la lógica epistémica?

Mi definición

Es la lógica de la representación del conocimiento y las creencias de un individuo (en I.A. un agente). A partir de la extensión de la lógica proposicional con operadores modales, podemos modelar de manera formal el poseer conocimientos y adquirirlos (en donde adquirir puede ser visto como una forma de razonamiento del individuo).

Super relacionada con Belief Revision (me parece bastante volátil el fin de una y el comienzo de otra).

Introduciendo a los nuevos operadores

En las lógicas epistémicas existen dos operadores modales. Si a representa un agente, escribimos:

- a conoce una fórmula ϕ como $K_a\phi$
- a cree una fórmula ϕ como $B_a\phi$

Para el problema que vamos a atacar, es indistinto el operador. En todo lo que sigue de la presentación, voy a usar el primer operador y lo voy a notar con la notación utilizada en toda la materia, pero debe entenderse que es indistinto usar cualquiera de los dos.

Queremos modelar a la epistemología, tenemos la sintaxis, los operadores ... Qué modelo semántico usamos?

Usando el modelo de Kripke

Nos encontramos con un problema conocido como el "logical omniscience problem". Sabemos que:

Lema

Si ϕ y $(\phi \implies \psi)$ son válidas en un modelo, también lo es ψ .

Y recordando que K tiene la regla de necesidad que dice:

Necesitación

Si ϕ es válida en un modelo, entonces también lo es $\Box\phi$.

Llegamos a que la lógica modal K es muy fuerte para modelar conocimiento y creencia!.

Neighbourhood al rescate!

Vamos a definir a una estructura epistémica, (los famosos modelos de vecindad con otro sabor), como una tripla $M = (W, N, I)$ en donde, si A es un conjunto de agentes, y $a \in A$ entonces:

- W es un conjunto de mundos
- $N : A \times W \mapsto 2^{2^W}$ es la función que asigna a cada agente en un mundo, el correspondiente conjunto de proposiciones que conoce (es decir, un conjunto epistémico).
- $I : P \mapsto 2^W$ es la función que asigna a cada proposición atómica, el conjunto de mundos en donde dicha proposición es satisfecha.

La definición de satisfacibilidad de una fórmula será la misma que vimos en la materia.

Problema con Neighbourhood

Si bien con el modelo definido recién, solucionamos el problema de logical omniscience, nos surge un nuevo problema. Citando la diapo 6 de la tercera clase:

Teorema

Si $\phi \iff \psi$ es válida en una clase de modelos de vecindad, entonces en dicha clase también lo es $\Box\phi \iff \Box\psi$.

Nos vamos a permitir vivir con este problema

El razonamiento como fórmulas

- 1 $\neg \Box \text{false}$
- 2 $\Box \text{true}$
- 3 $\Box(p \wedge q) \implies \Box q$
- 4 $(\Box p \wedge \Box q) \implies \Box(p \wedge q)$
- 5 $\Box p \implies \Box \Box p$
- 6 $\neg \Box p \implies \Box \neg \Box p$
- 7 $\Box p \implies p$

El razonamiento modelado sobre conjuntos epistémicos!

- 1 $\emptyset \notin N(a, w)$
- 2 $W \in N(a, w)$
- 3 Si $U \in N(a, w)$ y $U \subseteq V$, $\implies V \in N(a, w)$
- 4 Si $U \in N(a, w)$ y $V \in N(a, w) \implies (U \cap V) \in N(a, w)$
- 5 Si $U \in N(a, w) \implies \{u : U \in N(a, u)\} \in N(a, w)$
- 6 Si $U \notin N(a, w) \implies \{u : U \notin N(a, u)\} \in N(a, w)$
- 7 Si $U \in N(a, w) \implies w \in U$

A partir de esta definición, si llamamos ϵ a la clase de todos los modelos epistémicos, entonces podemos decir:

Definición

Sea S un subconjunto de $1, \dots, 7$, entonces ϵ_S será la clase de los modelos epistémicos que satisfaga $C_j \forall j \in S$

Complejidad computacional de cada modelo ?

Sabemos que una fórmula ϕ es ϵ_S -satisfacible si $\exists w \in W$ tal que $\epsilon_S \models_w \phi$. Cúal será la relación, si es que la hay, entre S y la complejidad computacional de satisfacer una fórmula?

EL problema a resolver en este paper

EL Teorema

Si S es un subconjunto de $1, \dots, 7$ entonces resolver SAT está en PSPACE.
Ahora, si S es un subconjunto de $1, \dots, 7$ y $4 \notin S$, entonces SAT está en NP

El paper de "On the Complexity of Epistemic Reasoning" demuestra este teorema (todo el paper está enfocado en esto).

La idea para resolver el problema

El objetivo para demostrar el teorema anterior viene por este lado

La idea

Sea M uno de los modelos epistémicos ϵ_i con $i \in S$, ϕ una fórmula tal que $\epsilon_i \models \phi$. La idea será probar que siempre que $i \neq 4$ existe una valuación v para ϕ tal que las subformulás con operadores modales en ϕ son satisfacibles, si ψ es satisfacible, con ψ una fórmula perteneciente al cálculo proposicional.

Con esto podemos aplicar el algoritmo de tableau de una manera más eficiente. Además, al dar propiedades de esta pinta, nos podemos construir un algoritmo no determinístico de tiempo polinomial.

Veamos un par de ejemplos de esta idea que comente

Proposición 1

Una fórmula ϕ es ϵ -satisfacible *iff* \exists v valuación para ϕ tal que si $\Box\psi_1$ y $\Box\psi_2 \in \text{sub}(\phi)$, $v(\Box\psi_1) = 1$, y $v(\Box\psi_2) = 0 \implies (\psi_1 \wedge \psi_2) \vee (\neg\psi_1 \wedge \psi_2)$ es ϵ -satisfacible

Proposición 2

Una fórmula ϕ es ϵ_2 -satisfacible *iff* \exists v valuación para ϕ tal que:

- Si $\Box\psi \in \text{sub}(\phi)$ y $sv(\Box\psi) = 0 \implies \neg\psi$ es ϵ_2 -satisfacible, y
- $\Box\psi_1$ y $\Box\psi_2 \in \text{sub}(\phi)$, $v(\Box\psi_1) = 1$, y $v(\Box\psi_2) = 0 \implies (\psi_1 \wedge \psi_2) \vee (\neg\psi_1 \wedge \psi_2)$ es ϵ_2 -satisfacible

Preguntas??