Operadores de Revisión Moderada

Daniel Grimaldi M. Vanina Martínez Ricardo O. Rodríguez

UBA, FCEyN, Departamento de Computación
UBA-CONICET, Inst. de Invest. en Cs. de la Computación

21 de septiembre de 2022

Introducción al problema de Cambio de Creencias

La representación del conocimiento se puede realizar sobre una lógica $\mathcal{L}$:

Introducción al problema de Cambio de Creencias

La representación del conocimiento se puede realizar sobre una lógica $\mathcal{L}$:

- K es un conjunto de fórmulas cerrado por consecuencia. Si $\mathcal{L}$ es finito, usamos una fórmula ψ. *Representa el conocimiento de une agente.*
- μ es una fórmula de $\mathcal{L}$. *Representa una observación del mundo real.*
- En nuestro contexto, necesitamos que ambos conceptos estén representados de la misma manera, por eso suponemos $\mathcal{L}$ finito.

Introducción al problema de Cambio de Creencias

La representación del conocimiento se puede realizar sobre una lógica $\mathcal{L}$:

- K es un conjunto de fórmulas cerrado por consecuencia. Si $\mathcal{L}$ es finito, usamos una fórmula ψ. *Representa el conocimiento de une agente.*
- μ es una fórmula de $\mathcal{L}$. *Representa una observación del mundo real.*
- En nuestro contexto, necesitamos que ambos conceptos estén representados de la misma manera, por eso suponemos $\mathcal{L}$ finito.

El problema: Adaptar un cuerpo de conocimiento a nueva información

Tenemos un cuerpo de conocimiento ψ de la realidad, y aparece una nueva información μ ¿Qué adaptación podríamos hacer sobre ψ?

Ejemplo concreto

Ejemplo

Hace unos meses mi pareja me dijo que no me ama más y nos separamos. Ayer gané la lotería. Hoy me llamó diciendo que se arrepentía de habernos separado y que me amaba.

Ejemplo concreto

Ejemplo

Hace unos meses mi pareja me dijo que no me ama más y nos separamos. Ayer gané la lotería. Hoy me llamó diciendo que se arrepentía de habernos separado y que me amaba.

$$p = \text{“Mi ex se arrepiente de la separación”}$$
$$q = \text{“Mi ex me ama”}$$
$$\psi = \neg q$$
$$\mu = p \wedge q$$

Ejemplo concreto

Ejemplo

Hace unos meses mi pareja me dijo que no me ama más y nos separamos. Ayer gané la lotería. Hoy me llamó diciendo que se arrepentía de habernos separado y que me amaba.

$$p = \text{“Mi ex se arrepiente de la separación”}$$
$$q = \text{“Mi ex me ama”}$$
$$\psi = \neg q$$
$$\mu = p \wedge q$$

En la literatura existen operadores conocidos, por ejemplo:

Revisar: $\psi * \mu = p \wedge q$ (se deduce μ)

Contraer: $\psi - (\neg\mu) = q \to p$ (no contradice μ)

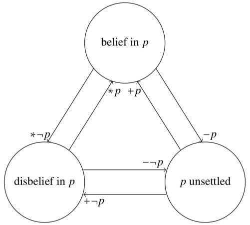
Si es verosímil opero, si no lo rechazo:

$$\psi \circ \mu = \begin{cases} \psi * \mu & \text{si } \mu \in \mathcal{C}(\psi) \\ \psi & \text{caso contrario} \end{cases}$$

Ideas conceptuales

En términos de estados epistémicos, con la revisión le agente cree en la nueva información. Con la contracción la nueva información queda incierta. En palabras más simples:

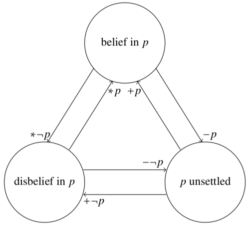
- Revisión es un operador de *certeza*.
- Contracción es un operador de *duda*.



Ideas conceptuales

En términos de estados epistémicos, con la revisión le agente cree en la nueva información. Con la contracción la nueva información queda incierta. En palabras más simples:

- Revisión es un operador de *certeza*.
- Contracción es un operador de *duda*.



Identidades entre Revision y Contraction

Identidad de Levi: $\psi * \mu \equiv (\psi - \neg\mu) \wedge \mu$
Identidad de Harper: $\psi - \mu \equiv (\psi * \neg\mu) \vee \psi$

- $\wedge$ lleva la contracción (duda) a la revisión (certeza).
- $\vee$ lleva la revisión (certeza) a la contracción (duda).

Un equilibrio entre revisión y contracción

Ejemplo

Hace unos meses mi pareja me dijo que no me ama más y nos separamos. Ayer gané la lotería. Hoy me llamó diciendo que se arrepentía de habernos separado y que me amaba.

Recuperando el ejemplo, nuestro agente, tiene varias opciones:

Un equilibrio entre revisión y contracción

Ejemplo

Hace unos meses mi pareja me dijo que no me ama más y nos separamos. Ayer gané la lotería. Hoy me llamó diciendo que se arrepentía de habernos separado y que me amaba.

Recuperando el ejemplo, nuestro agente, tiene varias opciones:

- Revisar, y creerle todo a su ex $(p \wedge q)$.
- Rechazar, y hacer como si no hubiera escuchado nada $(\neg q)$.
- Contraer, y creer que lo único claro es que si su ex le ama, entonces se arrepiente de la separación $(\neg q \vee (p \wedge q) \equiv q \rightarrow p)$.

Un equilibrio entre revisión y contracción

Ejemplo

Hace unos meses mi pareja me dijo que no me ama más y nos separamos. Ayer gané la lotería. Hoy me llamó diciendo que se arrepentía de habernos separado y que me amaba.

Recuperando el ejemplo, nuestro agente, tiene varias opciones:

- Revisar, y creerle todo a su ex $(p \wedge q)$.
- Rechazar, y hacer como si no hubiera escuchado nada $(\neg q)$.
- Contraer, y creer que lo único claro es que si su ex le ama, entonces se arrepiente de la separación $(\neg q \vee (p \wedge q) \equiv q \rightarrow p)$.

Una reacción más *natural* tal vez podría ser:

"*Puedo creer que quiera volver, pero dudo que me ame*" (p)

Un equilibrio entre revisión y contracción

Ejemplo

Hace unos meses mi pareja me dijo que no me ama más y nos separamos. Ayer gané la lotería. Hoy me llamó diciendo que se arrepentía de habernos separado y que me amaba.

Recuperando el ejemplo, nuestro agente, tiene varias opciones:

- Revisar, y creerle todo a su ex $(p \wedge q)$.
- Rechazar, y hacer como si no hubiera escuchado nada $(\neg q)$.
- Contraer, y creer que lo único claro es que si su ex le ama, entonces se arrepiente de la separación $(\neg q \vee (p \wedge q) \equiv q \rightarrow p)$.

Una reacción más *natural* tal vez podría ser:

"*Puedo creer que quiera volver, pero dudo que me ame*" (p)

Notar que: $$\psi * \mu \models p \models \psi - (\neg \mu)$$

Lo que buscamos está entre la certeza de la revisión y la duda de la contracción, una *revisión moderada*

Operadores en el Entorno de Katsuno-Mendelson

La adaptación del operador de revisión al caso finito (Katsuno and Mendelson, 1990) se define axiomáticamente por:

R 1. $\psi * \mu \models \mu$.

R 2. Si $\psi \wedge \mu \not\models \perp$ entonces $\psi * \mu \equiv \psi \wedge \mu$.

R 3. Si $\mu \not\models \perp$ entonces $\psi * \mu \not\models \perp$.

R 4. Si $\psi_1 \equiv \psi_2$ y $\mu_1 \equiv \mu_2$ entonces $\psi_1 * \mu_1 \equiv \psi_2 * \mu_2$.

Credibility Limited Revision (Hansson et. al. 2001) se define como:

$$\psi \circ \mu = \begin{cases} \psi * \mu & \text{si } \mu \in \mathcal{C}(\psi) \\ \psi & \text{caso contrario} \end{cases}$$

Y axiomáticamente por:

- **1.** $\psi \circ \mu = \psi$ o bien $\psi \circ \mu \models \mu$.
- **2.** $\psi \wedge \mu \models \psi \circ \mu$.
- **3.** Si $\psi \wedge (\psi \circ \mu) \not\models \perp$ entonces $\psi \circ \mu \models \psi$.
- **4.** Si $\psi \not\models \perp$ y $\mu \not\models \perp$ entonces $\psi \circ \mu \not\models \perp$.
- **5.** Si $\psi \circ \mu \models \mu$, $\phi \circ \nu \models \nu$, $\psi \equiv \phi$ y $\mu \equiv \nu$ entonces $\psi \circ \mu \equiv \phi \circ \nu$.

Moderated Revision

Definición de Moderated Revision

Un operador Moderated Revision se define de la siguiente manera, inducido por una revisión $*$ y una CL revision $\circ$:

$$\psi \circledast \mu = (\psi * \mu) \vee ((\mu \circ \psi) \wedge \psi)$$

Moderated Revision es un operador que no sólo está entre una revisión y una contracción, sino que extiende completamente revisión y parcialmente contracción. Utiliza el mecanismo de decisión de CL para determinar el grado de certeza que tiene el agente:

- Si $\psi \in \mathcal{C}(\mu)$ significa que el agente tiene ciertas dudas de μ.
- Si $\psi \notin \mathcal{C}(\mu)$ significa que el agente cree que μ es cierto.

Propiedades Elementales de Moderated Revision

Moderated Revision es un operador capaz de moverse por todos los estados epistémicos.

- $\psi * \mu \models \psi \circledast \mu \models \psi - \neg\mu$.
- Si $\psi \wedge \mu \not\models \bot$ entonces $\psi \circledast \mu \equiv \psi \wedge \mu$.
- Si $\mu \circ \psi \equiv \mu$ (i.e. $\psi \notin \mathcal{C}(\mu)$) entonces $\psi \circledast \mu \equiv \psi * \mu$.
- Si $\psi \wedge \mu \models \bot$ y $\mu \circ \psi \equiv \psi$ entonces $\psi \circledast \mu \equiv \psi - \neg\mu$.

Identidades con respecto a sus operadores que lo inducen y sus duales:

- $\psi \circledast \mu \equiv (\psi - \neg\mu) \wedge (\mu \div \neg\psi)$.
- $\psi * \mu \equiv (\psi \circledast \mu) \wedge \mu$.
- $\psi - \mu \equiv (\psi \circledast \neg\mu) \vee \psi$.
- $\mu \circ \psi \equiv \begin{cases} (\psi \circledast \mu) \wedge \psi & \text{si } (\psi \circledast \mu) \wedge \psi \not\models \bot \\ \mu & \text{caso contrario} \end{cases}$
- $\mu \div \psi \equiv (\neg\psi \circledast \mu) \vee \mu$.

Postulados de Moderated Revision

Teorema de Postulados de Moderated Revision

Las siguientes condiciones son equivalentes:

1. $\circledast$ **1.** $\psi \circledast \mu \models \psi \vee \mu$. (shared success)
   $\circledast$ **2.** Si $\psi \wedge \mu \not\models \bot$ entonces $\psi \circledast \mu \equiv \psi \wedge \mu$. (vacuity)
   $\circledast$ **3.** Si $\mu \not\models \bot$ entonces $(\psi \circledast \mu) \wedge \mu \not\models \bot$. (R-consistency)
   $\circledast$ **4.** Si $\psi \equiv \phi$ y $\mu \equiv \nu$ entonces $(\psi \circledast \mu) \wedge \mu \equiv (\phi \circledast \nu) \wedge \nu$. (R-extensionality)
   $\circledast$ **5.** Si $(\psi \circledast \mu) \wedge \psi \not\models \bot$, $(\phi \circledast \nu) \wedge \phi \not\models \bot$, $\psi \equiv \phi$ y $\mu \equiv \nu$ entonces $(\psi \circledast \mu) \wedge \psi \equiv (\phi \circledast \nu) \wedge \phi$. (CL-relative extensionality)
2. $\circledast$ es un operador moderated revision inducido por un operador de revision $*$ que satisface (R1) $\sim$ (R4) y una CL revision operator $\circ$.

Postulados de Moderated Revision

Teorema de Postulados de Moderated Revision

Las siguientes condiciones son equivalentes:

1. $\circledast$ **1.** $\psi \circledast \mu \models \psi \vee \mu$. (shared success)
   $\circledast$ **2.** Si $\psi \wedge \mu \not\models \bot$ entonces $\psi \circledast \mu \equiv \psi \wedge \mu$. (vacuity)
   $\circledast$ **3.** Si $\mu \not\models \bot$ entonces $(\psi \circledast \mu) \wedge \mu \not\models \bot$. (R-consistency)
   $\circledast$ **4.** Si $\psi \equiv \phi$ y $\mu \equiv \nu$ entonces $(\psi \circledast \mu) \wedge \mu \equiv (\phi \circledast \nu) \wedge \nu$. (R-extensionality)
   $\circledast$ **5.** Si $(\psi \circledast \mu) \wedge \psi \not\models \bot$, $(\phi \circledast \nu) \wedge \phi \not\models \bot$, $\psi \equiv \phi$ y $\mu \equiv \nu$ entonces $(\psi \circledast \mu) \wedge \psi \equiv (\phi \circledast \nu) \wedge \phi$. (CL-relative extensionality)
2. $\circledast$ es un operador moderated revision inducido por un operador de revision $*$ que satisface (R1) $\sim$ (R4) y una CL revision operator $\circ$.

Además, se tiene que:

$$\psi \in \mathcal{C}(\mu) \quad \Leftrightarrow \quad (\psi \circledast \mu) \wedge \psi \not\models \bot$$

Clasificación de Familias de Moderated Revision

Se identificaron tres familias “puras” basadas en las propiedades extras:

- conocidas de la revision $*$.
- conocidas sobre la imagen de $\mathcal{C}$ y de $*'$ (de la $\circ$).
- nuevas del dominio de $\mathcal{C}$ (de la $\circ$).

Puede haber “híbridos”, ya que las propiedades son complementarias.

Clasificación de Familias de Moderated Revision

Se identificaron tres familias "puras" basadas en las propiedades extras:

- conocidas de la revision $*$.
- conocidas sobre la imagen de $\mathcal{C}$ y de $*'$ (de la $\circ$).
- nuevas del dominio de $\mathcal{C}$ (de la $\circ$).

Puede haber "híbridos", ya que las propiedades son complementarias.

Propiedades extras de la revision $*$: Postulado de cambio mínimo:

$$\text{Si} \quad (\psi * \mu) \land \nu \not\models \bot \quad \text{entonces} \quad (\psi * \mu) \land \nu \equiv (\psi * (\mu \land \nu)) \land \nu$$

Clasificación de Familias de Moderated Revision

Se identificaron tres familias “puras” basadas en las propiedades extras:

- conocidas de la revision $*$.
- conocidas sobre la imagen de $\mathcal{C}$ y de $*'$ (de la $\circ$).
- nuevas del dominio de $\mathcal{C}$ (de la $\circ$).

Puede haber “híbridos”, ya que las propiedades son complementarias.

Propiedades extras de la revision $*$: Postulado de cambio mínimo:

$$\text{Si } \ (\psi * \mu) \wedge \nu \not\models \perp \ \text{ entonces } \ (\psi * \mu) \wedge \nu \equiv (\psi * (\mu \wedge \nu)) \wedge \nu$$

Propiedades extras sobre la imagen de $\mathcal{C}$ y de $*'$ (de la $\circ$).

- Puede existir $\mathcal{A}(\mu)$ tal que $\psi \in \mathcal{C}(\mu) \Leftrightarrow \psi \wedge \mathcal{A}(\mu) \not\models \perp$.
  En este contexto, además puede cumpir:
  - que si $\psi \circ \mu \not\models \perp$ entonces $(\psi \circ \mu) \wedge \mathcal{A}(\psi) \not\models \perp$.
  - que $\psi \circ \mu \models \mathcal{A}(\psi)$.
  - que el operador $*'$ satisfaga el postulado de cambio mínimo.

Clasificación de Familias de Moderated Revision

Propiedades extras sobre el dominio de $\mathcal{C}$.

- Puede haber irrelevancia de sintaxis: Si $\mu \equiv \nu$ entonces $\mathcal{C}(\mu) = \mathcal{C}(\nu)$.

Clasificación de Familias de Moderated Revision

Propiedades extras sobre el dominio de $\mathcal{C}$.

- Puede haber irrelevancia de sintaxis: Si $\mu \equiv \nu$ entonces $\mathcal{C}(\mu) = \mathcal{C}(\nu)$.
- Puede haber propiedades de un agente "imprudente":
  - Si $\mu \models \nu$ entonces $\mathcal{C}(\mu) \subset \mathcal{C}(\nu)$.
  - $\mathcal{C}(\mu \wedge \nu) = \mathcal{C}(\mu) \cap \mathcal{C}(\nu)$.
  - $\mathcal{C}(\neg\mu) \cap \mathcal{C}(\mu) = \emptyset$.
- Puede haber propiedades de un agente "sabio":
  - Si $\mu \models \nu$ entonces $\mathcal{C}(\nu) \subset \mathcal{C}(\mu)$.
  - $\mathcal{C}(\mu \wedge \nu) = \mathcal{C}(\mu) \cup \mathcal{C}(\nu)$.
  - $\mathcal{C}(\neg\mu) \cup \mathcal{C}(\mu) = \mathcal{L}$.

Clasificación de Familias de Moderated Revision

Propiedades extras sobre el dominio de $\mathcal{C}$.

- Puede haber irrelevancia de sintaxis: Si $\mu \equiv \nu$ entonces $\mathcal{C}(\mu) = \mathcal{C}(\nu)$.
- Puede haber propiedades de un agente "imprudente":
  - Si $\mu \models \nu$ entonces $\mathcal{C}(\mu) \subset \mathcal{C}(\nu)$.
  - $\mathcal{C}(\mu \wedge \nu) = \mathcal{C}(\mu) \cap \mathcal{C}(\nu)$.
  - $\mathcal{C}(\neg\mu) \cap \mathcal{C}(\mu) = \emptyset$.
- Puede haber propiedades de un agente "sabio":
  - Si $\mu \models \nu$ entonces $\mathcal{C}(\nu) \subset \mathcal{C}(\mu)$.
  - $\mathcal{C}(\mu \wedge \nu) = \mathcal{C}(\mu) \cup \mathcal{C}(\nu)$.
  - $\mathcal{C}(\neg\mu) \cup \mathcal{C}(\mu) = \mathcal{L}$.

"The whole problem with the world is that fools and fanatics are always so certain of themselves, but wiser people are so full of doubts"
Bertrand Russell

¿Dudas? ¿Certezas? ¿Rechazos?

¿Dudas? ¿Certezas? ¿Rechazos?

¡Gracias!