

Computational Logic and Human Thinking: How to be artificially intelligent

Robert A. Kowalski

Department of Computing
Imperial College

Traduce: Jacinto Dávila

1 Introducción

La Lógica Computacional ha sido desarrollada en el campo de la Inteligencia Artificial en los últimos 50 años, en un intento por programar computadores que muestren niveles humanos de inteligencia. Está basada en la Lógica Simbólica, lenguaje cuyas oraciones están representadas por símbolos y cuyo razonamiento se realiza a través de la manipulación de símbolos, tal como en la resolución de ecuaciones en el álgebra. Sin embargo, los intentos de usar la Lógica Simbólica para resolver problemas prácticos por medio de computadores la han conducido a muchas simplificaciones y extensiones. La Lógica Computacional resultante no solamente es más poderosa en cuanto a su aplicación computacional, sino que también y más útil para los el propósito original de la lógica: mejorar el pensar humano.

La Lógica Tradicional, la Lógica Simbólica y la Lógica Computacional tienen todas que ver con la forma abstracta de las oraciones y con el cómo esa forma afecta la correctitud de los argumentos. Aún cuando la lógica tradicional se remonta a la época de Aristóteles, en el siglo IV A.C., la Lógica Simbólica se desarrolla por vez primera en el siglo XIX, con las formas matemáticas de la lógica desarrolladas por George Boole y Gottlob Frege. Se le extiende considerablemente en el siglo XX con el trabajo de Bertrand Russell, Alfred North Whitehead, Kurt Gödel y muchos otros, con su aplicación a los fundamentos de las Matemáticas.

La Lógica Computacional emerge en la segunda mitad del siglo XX, a partir de intentos por mecanizar la generación de pruebas en matemáticas, y fué luego extendida tanto para representar formas más generales de conocimiento, como para realizar estrategias más generales para resolver problemas. La variedad de Lógica Computacional que se presenta en este libro debe mucho a las contribuciones de John McCarthy y John Alan Robinson.

Los logros de la Lógica Simbólica en el siglo pasado han sido considerables. Pero han conducido a que la Lógica originaria se convierta en una rama la Matemática y haya perdido contacto con sus raíces en el razonamiento humano. La Lógica Computacional también emplea notación matemática, lo cual facilita su implementación computacional pero oscurece su relevancia para el pensar humano.

En este libro, intentaré mostrar que los beneficios prácticos de la Lógica Computacional no están limitados a la Matemática y a la Inteligencia Artificial, sino que pueden ser también disfrutados por gente ordinaria en su vida diaria, sin tener que usar la notación matemática. No obstante, estoy incluyendo varios capítulos adicionales, más técnicos, al final del libro, que bien pueden ser omitidos por el lector causal (pero ayudarán a los interesados en las técnicas).

Introducción: La relación entre la lógica y el pensamiento

La Lógica, en todas sus variedades, tiene que ver con formalizar las leyes del pensamiento. Junto a disciplinas asociadas, como los estudios de Leyes y Derecho y las Ciencias Gerenciales, la Lógica se concentra en la formulación de teorías *normativas*, que prescriben como debería razonar la gente. La Psicología Cognitiva también está interesada en el pensamiento, pero se concentra casi exclusivamente en teorías *descriptivas*, que estudian cómo, de hecho, piensa la gente la práctica, sea correcto o no. En general, esos dos tipos de teorías se han desarrollado por separado y mantienen muy poca relación entre ellos.

Introducción: La relación entre la lógica y el pensamiento

No obstante ese divorcio, en los últimos años, los psicólogos cognitivistas han desarrollado teorías del *Proceso Dual* (Dual Process Theories), que pueden ser entendidas como la combinación de teorías normativas y descriptivas. Vistas desde la perspectiva de las teorías de Proceso Dual, las teorías descriptivas tradicionales se enfocan en el *pensamiento intuitivo*, que es asociativo, automático, paralelo y subconciente. Las teorías normativas tradicionales, por otro lado, se enfocan en el *pensamiento deliberativo* (N.T. se debería decir racional, deliberativo en Español no tiene esa acepción), que está basado en reglas, requiere un esfuerzo, es serial y conciente. En este libro, argumentaré que la Lógica Computacional es una teoría de proceso dual, en la cual se combinan el pensamiento deliberativo y el intuitivo.

Introducción: La relación entre la lógica y el pensamiento

Pero la Lógica está interesada, no solamente con el pensar en abstracto, sino también con pensamientos representados en forma de oraciones y con pensar visto como la manipulación de oraciones para generar nuevos pensamientos. En la Lógica Computacional, estas manipulaciones lógicas de las oraciones tienen también una interpretación computacional. Vista de esta manera, la Lógica Computacional puede ser vista como una formalización del lenguaje del pensamiento humano.

Introducción: La Lógica Computacional y el Lenguaje del Pensamiento

Tal y como se le usa en la Inteligencia Artificial, la Lógica Computacional funciona primero y principalmente como el *lenguaje de pensamiento* de un agente inteligente. Se incluye en ella una *sintaxis* (o gramática) que determina la forma de los pensamientos del agente, una *semántica*, que determina el contenido (o significado) de esos pensamientos, y un *motor de inferencia* (o procedimiento de prueba), que genera (o deriva, o infiere) nuevos pensamientos como consecuencia de los existentes. En este caso, la Lógica Computacional puede ser considerada como un *lenguaje privado*, que representa las metas y creencias de un agente y le ayuda a regular su conducta. Este lenguaje privado es independiente de, y más fundamental que, los lenguajes ordinarios como el Inglés (y el Español).

Introducción: La Lógica Computacional y el Lenguaje del Pensamiento

Sin embargo, en los sistemas multi-agente de la Inteligencia Artificial, el lenguaje privado de un agente individual sirve también la función secundaria de representar el significado de sus comunicaciones con otros agentes. Esas comunicaciones se expresan en un *lenguaje* compartido y por tanto *público*, que puede ser diferente de los lenguajes privados de los agentes individuales. La tarea de un agente que quiere comunicarse es la de traducir los pensamientos desde su lenguaje privado hacia el lenguaje público, de forma tal que el agente receptor pueda traducir fácilmente esas comunicaciones públicas en los pensamientos apropiados representados en su propio lenguaje privado.

Sería más fácil si todos los agentes compartieran el mismo lenguaje privado, y si tal lenguaje fuese idéntico al lenguaje público de la comunidad de agentes. Esto se puede lograr al diseñar un sistema multi-agente artificial, pero sólo puede ser aproximado en una sociedad de agentes humanos.

Introducción: La Lógica Computacional y el Lenguaje del Pensamiento

La distinción entre lenguajes públicos y privados, tan clara en la Inteligencia Artificial, ha sido propuesta en el campo de la Filosofía del Lenguaje para explicar la relación entre el pensar humano y la comunicación. Muchas de esas propuestas, que por simplicidad pueden ser todas denominadas como propuestas de un Lenguaje del Pensamiento, LOT (de *Language of Thought*), sostienen que, en buena medida, el pensar humano se puede decir que ocurre en un lenguaje del pensamiento. La más famosa de esas propuestas es la hipótesis de Fodor de que el LOT es un lenguaje privado, independiente de la Babel de lenguajes públicos (Fodor, 1975). Otras propuestas, en particular (Carruthers, 2004) argumentan que el LOT de una persona es específico al lenguaje público de la comunidad social de esa persona.

Introducción: La Lógica Computacional y el Lenguaje del Pensamiento

Sin importar su posición acerca de la relación entre lenguajes privados y públicos, la mayoría de las propuestas parecen estar de acuerdo en que el LOT tiene cierto tipo de forma lógica. Sin embargo, la mayor parte de esas propuestas son particularmente tímidas al dar detalles de esa forma lógica. En comparación con aquellas, la propuesta que presento en este libro - que la Lógica Computacional puede ser considerada como una formalización de el LOT - es desvergonzadamente reveladora. Para mi argumento, invoco el apoyo de los usos que se da a la Lógica Computacional en la Inteligencia Artificial. Pero también invoco apoyo desde la relación entre la Lógica Computacional y las teorías normativas de la comunicación humana.

Introducción: Lógica Computacional y la Comunicación Humana

Normalmente, cuando hablamos o escribimos, simplemente nos expresamos en público sin hacer un esfuerzo consciente para comunicarnos efectivamente. Pero cuando realmente nos importa ser entendidos - como me ocurre al escribir este libro - tratamos de ser tan claros, coherentes y convincentes como sea posible. La diferencia es como la diferencia entre las teorías descriptivas y las normativas del pensamiento; y, como ocurre en los dos casos de pensamiento, las dos clases de comunicación se estudian mayormente en disciplinas académicas diferentes.

Introducción: Lógica Computacional y la Comunicación Humana

Mientras la Lingüística se ocupa de desarrollar teorías descriptivas acerca del cómo usa la gente el lenguaje en la práctica, la Retórica y sus disciplinas aliadas, tales como la Composición (Redacción) y el Pensamiento Crítico, se ocupan de teorías normativas acerca del cómo debería la gente usar el lenguaje para comunicarse más efectivamente.

En este libro, presento una teoría normativa sobre el pensar, la comunicación y la conducta inteligente. Pero le presto atención a las teorías descriptivas, porque ayudan a entender de donde venimos, mientras que la normativas nos muestran a donde debemos ir.

Introducción: Lógica Computacional y la Comunicación Humana

La teoría descriptiva de la comunicación que se acerca más a una teoría normativa es la Teoría de la Relevancia (Sperber y Wilson, 1986). Está basada en una teoría más general de la cognición, la cual, parafraseando superficialmente, propone la hipótesis de que, dados estímulos o entradas alternativas desde su ambiente, la gente presta atención a aquellas entradas que proporcionan la información más útil por el menor costo de procesamiento. Aplicadas a la comunicación, la teoría propone que, dada una comunicación potencialmente ambigua como entrada, los lectores o escuchas traducirán esa entrada en una forma lógica que maximiza la cantidad de información que contiene, mientras que minimiza el esfuerzo computacional necesario para generar la forma lógica.

Introducción: Lógica Computacional y la Comunicación Humana

La Teoría de la Relevancia es compatible con la hipótesis de que la Lógica Computacional, o algo parecido, es la lógica del lenguaje del pensamiento. Como la Lógica Computacional, la Teoría de la Relevancia tiene tanto componentes lógicos como computacionales. Más aún, proporciona una conexión con teorías normativas de la comunicación como las guías de Joseph Williams acerca de estilos de escritura en Inglés (Williams, 1990/1995).

Introducción: Lógica Computacional y la Comunicación Humana

Una forma de interpretar las recomendaciones de Williams y traduciéndolas a términos lógicos, lo cual implica el consejo de que los escritores deben expresarse en una forma que sea tan cercana como sea posible a la forma lógica de los pensamientos que ellos desean comunicar. En otras palabras, deben decir claramente lo que quieren decir y deben hacerlo en forma tal que sean tan fácil como sea posible de que los lectores extraigan ese significado. O, para decirlo todavía de otra manera, la expresión pública de nuestros pensamientos privados debe ser tan cercana como sea posible a la forma lógica de esos pensamientos.

Introducción: Lógica Computacional y la Comunicación Humana

Si nuestro lenguaje privado y lenguaje público fuesen el mismo, podríamos simplemente decir lo que pensamos. Pero incluso eso podría no ser suficiente. Todavía tendríamos que organizar nuestros pensamientos coherentemente, de forma que cada pensamiento estuviera lógicamente conectado con otro, y de forma tal que nuestros lectores o escuchas pudieran conectar nuestros pensamientos a los suyos.

La recomendación de Williams acerca del cómo obtener esa coherencia incluye el consejo de colocar ideas viejas o familiares al comienzo de una oración y colocar las nuevas ideas al final. En una serie de oraciones, una nueva idea al final de una oración se convierte en una vieja idea que puede colocarse al comienzo de la siguiente oración.

Introducción: Lógica Computacional y la Comunicación Humana

Considere este ejemplo de ese consejo, que una versión informal de la sintaxis de la Lógica Computacional y que, por cierto, muestra como la Lógica Computacional puede ser usada para representar las metas y creencias de un agente para guiar su conducta:

Tu quieres ser más inteligente.

Tu serás más inteligente si eres más lógico o lógica.

Serás más lógico o lógica si estudias este libro.

Así que (si no hay otras alternativas) tu deberías estudiar este libro.

Quizás no sea poético y puede que Ud no esté de acuerdo, pero cuando menos es claro, coherente y al grano.

Introducción ¿Qué es la Lógica Computacional?

La versión de la Lógica Computacional que se presenta en este libro combina una forma simplificada de lenguaje para representar información con formas mecánicas (o automáticas) de usar esa información para inferir sus consecuencias. Las oraciones en ese lenguaje tiene la forma simple de las condicionales: *si condiciones entonces conclusión* (o, en forma equivalente, *conclusión si condiciones*). Las reglas básicas de inferencia son el razonamiento hacia adelante y hacia atrás.

Introducción ¿Qué es la Lógica Computacional?

El *razonamiento hacia adelante* es la regla clásica de inferencia (también llamada *modus ponens*) para derivar *conclusiones* de *condiciones*. Por ejemplo, dada la creencia que, en general, *una persona será más lógica si esa persona estudia este libro*, el razonamiento hacia adelante deriva la conclusión de que *María será más lógica* a partir de la condición *María estudia este libro*. El razonamiento hacia adelante incluye el caso especial en cual una agente deriva las consecuencias de sus observaciones, para determinar cómo esas consecuencias podrían afectar sus metas.

Introducción ¿Qué es la Lógica Computacional?

El *razonamiento hacia atrás* funciona en la dirección opuesta, para derivar *condiciones* a partir de *conclusiones*. Por ejemplo, dada la creencia de que en general *una persona será más inteligente si esa persona es más lógica* como la *única* manera de concluir que una persona será más inteligente, el razonamiento hacia atrás deriva la condición de que *Juan debe ser más lógico* de la conclusión *Juan será más inteligente*. El razonamiento hacia atrás puede ser considerado una forma de reducción de metas, en la cual la *conclusión* es una meta y las *condiciones* son submetas. El razonamiento hacia atrás incluye el caso especial en el cual un agente deriva submetas que son acciones, que el agente puede realizar en el mundo.

Introducción ¿Qué es la Lógica Computacional?

El razonamiento hacia atrás le da a la Lógica Computacional el poder de un lenguaje de programación de alto nivel, en el cual todos los programas consisten de procedimientos de reducción de metas. De hecho, el lenguaje de programación Prolog, que proviene de Programando en Lógica, explota esa forma de computación fundamentalmente para aplicaciones en Inteligencia Artificial.

La Lógica Computacional, en la forma más general que investigamos en este libro, también incluye el uso de inferencia para ayudar a un agente a escoger entre cursos de acción alternativos

Introducción ¿Qué es la Lógica Computacional?

Por ejemplo, habiendo usado el razonamiento hacia atrás para derivar dos submetas alternativas, digamos que *Juan es más lógico* o *Juan toma drogas que aumentar su inteligencia*, para alcanzar la meta *Juan es más inteligente*, Juan puede usar razonamiento hacia adelante para inferir las posibles consecuencias de las alternativas antes de decidir qué hacer. En particular, si Juan infiere la consecuencia de que *Juan puede sufrir daño cerebral irreversible* si Juan elige la segunda alternativa, *Juan toma drogas para aumentar su inteligencia*, entonces esto debería animar a Juan a elegir la primera alternativa, *Juan es más lógico*.

Introducción ¿Qué es la Inteligencia Artificial?

La Inteligencia Artificial, IA, es el intento de programar computadores que se comporten inteligentemente, según los estándares humanos. Las aplicaciones de la IA incluyen áreas de problemas tales como el reconocimiento de la voz, sistemas expertos para diagnósticos de fallas en medicina e ingeniería, y la formalización del razonamiento legal.

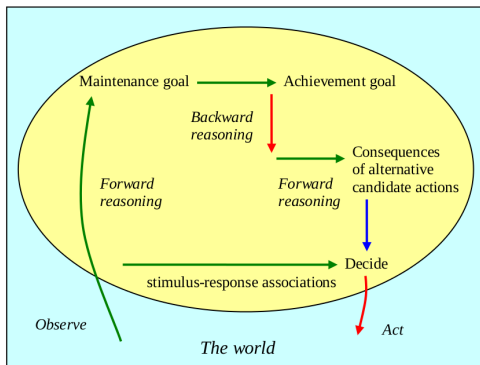
Las herramientas de la IA incluyen técnicas tales como búsqueda, Lógica Simbólica, redes neuronales artificiales y razonamiento con incertidumbre. Muchas de esas herramientas han contribuido al desarrollo de la Lógica Computacional que se investiga en este libro. Sin embargo, en lugar de preocuparnos con las aplicaciones de la IA, nos concentraremos en el uso de la Lógica Computacional para ayudar a la gente común a pensar y comportarse más inteligentemente.

Introducción ¿Qué es la Lógica Computacional?

Pensar en la gente meramente en términos computacionales podría sugerir que la gente puede ser tratada como si fuesen máquinas solamente. Por el contrario, creo que pensar en las personas como agentes computacionales puede ayudarnos a apreciar mejor nuestra naturaleza común y nuestras diferencias individuales. Eso destaca nuestra necesidad común de lidiar con el ciclo de la vida en un mundo siempre cambiante; y también llama la atención sobre el hecho de que otras personas pueden tener otras experiencias, metas y creencias, que son diferentes de las nuestras, pero que son igualmente dignas de comprensión, tolerancia y respeto.

Introducción: La Lógica Computacional y el Ciclo de la Vida

El papel de la Lógica Computacional en la mente de un agente inteligente puede ser dibujado aproximadamente así:



Introducción: La Lógica Computacional y el Ciclo de la Vida

Al ver la relación entre el agente y el mundo de esta manera, la mente del agente es una estructura sintáctica, que representa las creencias del agente acerca del mundo tal como lo ve y sus metas acerca del cómo le gustaría que ese mundo fuese. Estas creencias y metas son representadas en el lenguaje de pensamiento privado o propio de ese agente, cuyas oraciones tienen la forma sintáctica de las condicionales.

El mundo, por su lado, es una estructura semántica, que incluye el cuerpo del agente y le otorga significados a los pensamientos del agente. Es una estructura dinámica, que está constantemente cambiando y que existe sólo en el aquí y ahora. Sin embargo, el agente puede memorizar sus experiencias de ese cambio en su lenguaje de pensamiento, y puede formular creencias generales acerca de las relaciones causales entre sus experiencias. Puede entonces usar esas creencias, que explican sus experiencias pasadas, para ayudarse a alcanzar sus metas en el futuro.

Introducción: La Lógica Computacional y el Ciclo de la Vida

El agente observa eventos que tienen lugar en el mundo, así como las propiedades que esos eventos comienzan y terminan. Luego usa razonamiento hacia adelante para derivar conclusiones de sus observaciones. En muchos casos, esas conclusiones son acciones, disparadas por asociaciones estímulo-respuesta instintivas o intuitivas, que pueden también expresarse en la forma lógica de las condicionales. El agente puede ejecutar esas acciones por reflejo, automática e inmediatamente. Or, puede que les haga seguimiento, con una forma de razonamiento de alto nivel, como se postula en los modelos de proceso dual del pensamiento humano.

Introducción: La Lógica Computacional y el Ciclo de la Vida

Sin embargo, sea el agente tentado a reaccionar inmediatamente con asociaciones estímulo-respuesta o no, el agente puede también razonar hacia adelante para determinar si las observaciones afectan algunas metas de nivel superior que deben mantenerse para preservar una relación armoniosa con su ambiente. Razonar hacia adelante con metas de mantenimiento de alto nivel de este tipo genera metas por lograr en el futuro. El agente puede razonar hacia atrás, para reducir esas metas por lograr a submetas y buscar en su mente los planes de acción para alcanzar aquella metas.

Introducción: La Lógica Computacional y el Ciclo de la Vida

El agente puede entonces encontrar que hay varios planes alternativos para alcanzar la misma meta. Tendrá que decidir entre ellos. En la Teoría de Decisión clásica, el agente usa las consecuencias esperadas de su planes candidatos para ayudarse en tomar la decisión. Con sus creencias representadas en la forma lógica de los condicionales, esas consecuencias pueden ser derivadas al razonar hacia adelante a partir de las condiciones que representan el rendimiento hipotético las acciones candidatas alternativas. El agente puede evaluar las consecuencias, rechazar las acciones que tengan consecuencias no deseadas y escoger las acciones que tengan los resultados esperados (o la utilidad) más deseables.

Introducción: La Lógica Computacional y el Ciclo de la Vida

Sin embargo, las consecuencias de las acciones de un agente pueden depender, no sólo de sus propias acciones, sino también de acciones de otros agentes o de otras condiciones que están fuera del control del agente. El agente podría no ser capaz de determinar con exactitud si esas condiciones se cumplen por adelantado, pero podría juzgar su verosimilitud (o probabilidad). En tales casos, el agente puede usar técnicas de la Teoría de Decisiones, para combinar sus juicios de probabilidad y utilidad, y elegir el curso de acciones que tenga la mayor utilidad esperada.

Alternativamente, el agente puede usar planes de acción precompilados, más pragmáticos, que aproximen ese ideal de la Teoría de Decisión.

Introducción: La Lógica Computacional y el Ciclo de la Vida

Entre los criterios que un agente puede usar para decidir entre maneras alternativas de alcanzar sus metas, se cuenta el impacto probable en las metas de otros agentes. Las alternativas que ayuden a otros agentes a alcanzar sus metas, que no les limiten en el logro de sus metas, podrían recibir preferencia sobre otras alternativas. De esta forma, al ayudar al agente a entender y apreciar que otros agentes tienen sus propias experiencias, metas y creencias, la Lógica Computacional puede ayudar al agente a evitar conflictos y cooperar con otros agentes.

Este libro se orienta a mostrar que esos beneficios de la Lógica Computacional, que han tenido algún éxito en el campo de la Inteligencia Artificial, tienen también un gran potencial para mejorar el pensamiento y la conducta humanos.