

# Computational Logic and Human Thinking: How to be artificially intelligent

Robert A. Kowalski

Department of Computing  
Imperial College

Traduce: Jacinto Dávila

- 1 Presentación
- 2 Prefacio
- 3 Resumen del Libro

- 1 Presentación
- 2 Prefacio
- 3 Resumen del Libro

- 1 Presentación
- 2 Prefacio
- 3 Resumen del Libro

Los beneficios prácticos de la lógica computacional no tienen que estar limitados a las matemáticas y a la computación. Como muestra este libro, cualquier persona puede beneficiarse en su vida diaria de los últimos avances de la inteligencia artificial. Este libro se construye sobre desarrollos similares en varios campos, desde la filosofía, la psicología y hasta las leyes. El libro presta atención especial a la integración de la lógica con la teoría de decisiones, y en el uso de la lógica para mejorar la claridad y la coherencia de la comunicación en lenguajes naturales como el Inglés (y el Español).

Este libro es de lectura esencial para profesores e investigadores que quieran ponerse al día con los últimos desarrollos en lógica computacional. Puede también ser muy útil en cualquier curso de pregrado que enseñe pensamiento práctico, resolución de problemas o estrategias de comunicación. Su presentación informal lo convierte en un texto accesible para lectores con cualquier nivel de preparación previa. Sin embargo, el texto incluye también capítulos opcionales, pero más formales, para quienes tienen un propósito técnico en mente.

La mera posibilidad de la Inteligencia Artificial (IA) - la existencia de máquinas que puedan pensar y actuar con tanta inteligencia como los humanos - puede generar reacciones emotivas muy fuertes. Mientras los entusiastas se animan con la idea de que un día las máquinas sean más inteligentes que las personas, muchos críticos ven tal prospecto con horror.

En parte debido a que esta controversia atrae tanta atención, uno de los más importantes logros de la IA ha pasado sin ser notado: el hecho de que muchos de sus avances pueden también ser usados por las personas, para mejorar su propia inteligencia humana. Clave entre esos avances es la Lógica Computacional.



La Lógica Computacional se construye sobre la lógica tradicional, que fuera desarrollada originalmente para ayudar a la gente a pensar con mayor efectividad. La Lógica Computacional emplea las técnicas de la lógica simbólica, que han sido usadas para construir los fundamentos de la matemática y la computación. Sin embargo, comparada con la lógica tradicional, la Lógica Computacional es mucho más poderosa; y comparada con la lógica simbólica, es mucho más simple y más práctica.

Aun cuando las aplicaciones de la Lógica Computacional en IA requieren el uso de la notación matemática, sus aplicaciones humanas no.

Aprovechando esa circunstancia, he escrito la parte principal de este libro informalmente, para que llegue a una audiencia tan amplia como sea posible. Puesto que el pensamiento humano es también el objeto de estudio de muchos otros campos, he tomado prestado de estudios relacionados en Psicología Cognitiva, Lingüística, Filosofía, Derecho, Ciencias Gerenciales y Composición de textos en Inglés.

De hecho, la variante de la Lógica Computacional que se presenta en este libro se construye, no sólo sobre los desarrollos lógicos en IA, sino también sobre muchos otros paradigmas complementarios y alternativos para representación del conocimiento y resolución de problemas. En particular, incorpora representaciones procedimentales del conocimiento tal como se usan en la IA y la computación, sistemas de producción como en la IA y en la Ciencia Cognitiva, y el análisis de decisiones de las Ciencias Gerenciales, la Psicología Cognitiva y la Filosofía.

Debido a que la Lógica Computacional tiene tantas aplicaciones y tantas relaciones con otros campos, el uso ideal para este libro sería como texto de acompañamiento en un curso de pregrado en pensamiento práctico. Tal curso de pregrado combinaría las virtudes tradicionales de una educación liberal en las artes con las habilidades de argumentación de la filosofía analítica, el rigor del método científico y los beneficios modernos de la tecnología de la información. Le proveería al o a la estudiante con el pensamiento transferible y las habilidades de comunicación necesarias, no solamente para estudios más especializados, sino también para resolver problemas que no coinciden claramente con las áreas clasificadas.

Hasta donde sé, no existe hoy en día nada parecido a tal curso y, hasta donde puedo ver, es poco probable que exista en el futuro cercano. La Lógica, como la disciplina académica tal como existe hoy, está fragmentada y dispersa entre de la Matemática, la Filosofía y la Computación. Peor aún, las aplicaciones prácticas de la Lógica Informal están mayormente sepultadas en otras disciplinas académicas como el Derecho, la Ciencia Gerencial y la Composición Lingüística. Ninguna de esas disciplinas podría acomodar, por sí sola, tal curso, y pocas de ellas recibirían con agrado una expansión de la Lógica dentro de su propio campo.

Quizás un día, alguna institución educativa hará espacio para tal curso sobre el cómo pensar. Entretanto, este libro puede ser usado como complemento para cursos más convencionales. Y para quienes ya han completado su educación formal, el libro puede darles una visión de un mundo posible.

Al escribir este libro, he hecho grandes esfuerzos para evitar presentar mal el tema debido a una exagerada simplificación. Por tanto, he incluido una serie de capítulos opcionales, en los cuáles se encuentran los detalles técnicos que se evitan en los otros capítulos. Todos esos opcionales fueron escritos luego de completar el texto principal y se los puede saltar en una primera lectura u omitir completamente. Considerados por separado, proporcionan una introducción auto-contenida y las referencias a los soporte formales de la Lógica Computacional.

También he tenido presente el hecho de que, debido a las observaciones que hago acerca de estilos de escritura en Inglés, me expongo a la crítica de mi propio estilo al escribir. En mi defensa, permítanme decir que sin la ayuda de la Lógica Computacional, mi escritura sería mucho peor.

Cuando comencé mis estudios de pregrado en la Universidad de Chicago, hace muchos años, mi escritura era tan mala que reprobé el examen de ingreso y tuve que tomar un curso de reparación sin créditos. Terminé aquel año con A en todos mis cursos, pero con una D en escritura en Inglés. Me tomó años diagnosticar el problema en mi escritura y aprender a mejorarla. En el proceso de hacerlo, aprendí más acerca de la lógica práctica que lo que aprendí en mis cursos formales de lógica. Me gusta pensar que mi escritura es mucho mejor ahora que durante mi primer año en Chicago. Pero lo más importante es que confío que las lecciones que aprendí serán útiles a algunos lectores de este libro.

Dicté un curso corto, basado en este libro, en el Centro Internacional de Lógica Computacional (y con una versión previa en la Universidad de Los Andes en Venezuela).



# Prefacio y Agradecimientos

Estoy muy agradecido con Tom Blackson, François Bry, Tony Burton, Keith Clark, Jacinto Dávila, Phan Minh Dung, Maarten van Emden, Steffen Hoelldobler, Luis Pereira, Yongyuth Permpoontanalarp, Fariba Sadri, John Stenning, Dania Kowalska-Taylor, Sten-Ake Tarnlund, Jeff Thompson, Francesca Toni and Mike Tunstall por sus valiosos comentarios en los borradores previos de este libro. Agradezco también a Simon Taylor por los dibujos de la zorra y el cuervo

# Resumen y Plan del Libro

En virtud de que este libro recorre una gran variedad de tópicos, es útil resumir las relaciones entre los capítulos en un mismo lugar. Sin embargo, en lugar de colocar este resumen al final del libro, para cuando todo sus términos hayan sido explicados en detalle, he decidido presentarlo aquí, para ajustarme al espíritu general del libro que dice que mejor trabajar hacia atrás desde el destino, que tropezar hacia adelante, preguntándose hacia donde uno va. En consecuencia, este resumen puede ser leído tanto antes como después del cuerpo principal del libro. Pero también puede ser leído en paralelo al texto, para obtener una mejor orientación acerca del cómo se relacionan los diferentes capítulos.

# Resumen de la Introducción

En Inteligencia Artificial, un agente es cualquier entidad, imbuida en un mundo real o artificial, que puede observar el mundo cambiante y realizar acciones sobre ese mundo para mantenerse en una relación armoniosa con él. La Lógica Computacional, tal como se le usa en la Inteligencia Artificial, es el lenguaje de pensamiento del agente. Las oraciones que se expresan en ese lenguaje representan las creencias del agente acerca del cómo es el mundo y sus metas acerca del cómo le gustaría que fuera. El agente usa sus metas y creencias para controlar su conducta.

El agente usa las reglas de inferencia de la Lógica Computacional, aplicándolas a sus pensamientos en forma lógica, para razonar acerca del mundo y calcular sus acciones para cambiar el mundo para su propio beneficio. Tales reglas de inferencia incluyen tanto las de razonamiento hacia adelante, para derivar consecuencias a partir de sus observaciones, como las de razonamiento hacia atrás, para reducir sus metas en submetas y acciones. El agente puede también usar el razonamiento hacia adelante para deducir consecuencias de sus candidatos a acciones y así ayudarse a escoger entre las alternativas.

Aún cuando el propósito principal de la Lógica Computacional es representar los pensamientos privados de un agente y para controlar su conducta, el agente también puede usarla para guiar sus comunicaciones públicas con otros agentes. Al expresar sus comunicaciones en una forma más lógica, un hablante o escritor puede hacer que resulte más fácil para el escucha o el lector traducir esas comunicaciones en sus propios pensamientos.

# Resumen del Capítulo 1: La Lógica del Metro

El aviso de emergencia en metro de Londres ilustra cómo los significados de las comunicaciones en Inglés (y en Español) pueden comprenderse como pensamientos en forma lógica. En la Lógica Computacional, esos pensamientos pueden tener tanto un carácter lógico como uno computacional. Su carácter lógico es evidente en el uso explícito de las conectivas lógicas, como cualquier, si, y y no; y su carácter computacional se manifiesta en su uso como procedimientos para reducir metas en submetas. Debido a este carácter dual lógico y computacional, las oraciones que se expresan de esa manera son llamadas también programas lógicos.

# Resumen del Capítulo 1: La Lógica del Metro

El Aviso de emergencia también muestra como el uso coherente de las comunicaciones en Inglés (y en Español) puede ser entendido en términos de las conexiones lógicas entre el significado de esas conexiones y otros pensamientos en la red de metas y creencias de un agente. Una vez que el agente ha hecho las conexiones, el agente puede activarlas razonando hacia adelante y hace atrás, cuando hace falta. Las conexiones que son activadas con mayor frecuencia pueden ser convertidas en metas y creencias pre-derivadas, las cuáles pueden ser usadas más directamente y más eficientemente, más adelante.

## Resumen del Capítulo 2: La Psicología de la Lógica

El más influyente y más ampliamente citado argumento en contra de la lógica viene de los experimentos psicológicos acerca del razonar con oraciones en lenguaje natural y en forma condicional. La interpretación más popular de esos experimentos reza que las personas no tienen una habilidad natural y de múltiple propósito para razonar lógicamente, sino que en su lugar han desarrollado, a través de los mecanismos de la evolución darwiniana, algoritmos especializados para resolver problemas típicos que surgen en su ambiente.



## Resumen del Capítulo 2: La Psicología de la Lógica

En este capítulo, discuto algunos de aspectos que tienen que ver con la resolución de esas tareas de razonamiento y argumento que uno de los problemas principales con esos experimentos es que no alcanzan a apreciar que la forma en lenguaje natural de una oración condicional es sólo una aproximación a la forma lógica de su significado pretendido. Otro problema es que la interpretación de esos experimentos está basada en una comprensión inadecuada de la relación entre conocimiento y razonamiento. En la Lógica Computacional aplicada al pensamiento humano, esta relación puede ser expresada aproximadamente con una ecuación: Pensamiento = conocimiento especializado + razonamiento de propósito general.

## Resumen del Capítulo 3: La Zorra y el Cuervo

La Fábula de Esopo de la Zorra y el Cuervo ilustra el razonar hacia atrás de una muy astuta zorra, que genera un plan para alcanzar la meta de obtener el queso que sostiene un no muy astuto cuervo. Se establece un contraste entre el razonamiento proactivo y hacia atrás, y el razonamiento reactivo y hacia adelante del cuervo, quien responde a las alabanzas de la zorra iniciando el canto, con lo cual deja caer el queso al piso, desde donde la zorra puede recogerlo.

## Resumen del Capítulo 3: La Zorra y el Cuervo

Tanto la zorra como el cuervo razonan de acuerdo con las reglas de inferencia de la Lógica Computacional, pero la zorra tiene mejor conocimiento del mundo y formas más poderosas de usar ese conocimiento en su propio beneficio (N.T: Sin mencionar muy escasos escrúpulos). Si el cuervo supiera tanto como la zorra y fuese capaz de razonar preactivamente, pensando antes de actuar, podría entonces razonar hacia adelante a partir de la hipotética ejecución de sus candidatas a acción, predecir sus probables consecuencias y escoger una acción alternativa, como salir volando o tragarse el queso, que le permitiría alcanzar un mucho mejor estado final para sí mismo.

## Resumen del Capítulo 4: Búsqueda

En la Lógica Computacional, un procedimiento de prueba consiste de una colección de reglas de inferencias y una estrategia de búsqueda. Las reglas de inferencia determinan tanto la estructura de las pruebas como el espacio de búsqueda de todas las posibles pruebas que sean relevantes a la solución de una meta. La estrategia de búsqueda determina la manera en la cual el espacio de búsqueda es revisado mientras se busca una solución. Diversas estrategias de búsqueda son posibles, incluyendo tanto las estrategias paralelas, que exploran diferentes partes de espacio de búsqueda al mismo tiempo, como las estrategias de primero el mejor, que apuntan a encontrar la mejor solución posible en el tiempo más breve.

# Resumen del Capítulo 5: Negación por Falla

En la semántica de la Lógica Computacional, el mundo es un lugar positivo, caracterizado por las oraciones atómicas positivas que son ciertas en el momento. Puesto que el propósito último de las creencias y metas de un agente es gestionar sus interacciones con el mundo, las formas sintácticas de los pensamientos del agente también tienen un sesgo positivo correspondiente. En muchos casos, pensamientos sintácticamente negativos surgen de la falla al observar o derivar información positiva.

## Resumen del Capítulo 5: Negación por Falla

La negación por falla es una manera natural de razonar por omisión con información incompleta, derivando conclusiones a partir de la suposición implícita de que el agente lo sabe todo, pero luego retractándose graciosamente de esas conclusiones si alguna nueva información prueba que no se cumplen. La negación por falla también facilita formas superiores de organizar metas y creencias en jerarquías de reglas y excepciones, en las cuáles las reglas representan solamente la condición más importantes y las excepciones agregan condiciones adicionales cuando hacen faltan.

# Resumen del Capítulo 6: Cómo convertirse en ciudadano británico

El Acta de nacionalidad británica es una colección de oraciones en inglés que establece, con precisión, la condiciones bajo las cuáles una persona puede adquirir, renunciar o ser privada de la ciudadanía británica. El acta está diseñada para ser tanto inambigua, de manera que haya poca duda acerca del significado pretendido, como flexible, de manera que pueda ser aplicada en circunstancias cambiantes. Su estilo de Inglés se asemeja a la forma condicional de las oraciones en la Lógica Computacional.

# Resumen del Capítulo 6: Cómo convertirse en ciudadano británico

Además de ese uso la forma condicional, el acta de nacionalidad británica ilustra muchas otras características importantes acerca de la Lógica Computacional, incluyendo la representación de reglas y excepciones, razonamiento al meta-nivel acerca de lo que hace falta para que una persona, como Ud o como yo, satisfaga al Secretario de Estado en cuanto a que la persona cumple los requisitos para naturalizarse como ciudadano británico. Contrastando con el acta de nacionalidad británica, la cláusula de terminación de contrato de la Universidad de Michigan demuestra como un texto ambigüo, virtualmente inentelible, puede ser comprendido al reformularlo en la Lógica Computacional.



# Resumen del Capítulo 7: El Piojo y el Explorador de Marte

Se puede argumentar que el más influyente modelo computacional del pensamiento humano en psicología cognitivista es el modelo de los sistemas de producción, tal como se ilustra en este capítulo con la termita (N.T. el insecto que al cual el texto original parece referirse es la termita, aunque el término en inglés se refiere a la cochinillas o bichos bolitas que también comen madera) y el robot explorador de Marte. Los sistemas de producción combinan una memoria de trabajo de hechos atómicos con reglas de condición-acción con la forma si condiciones entonces acciones. La memoria de trabajo es como un modelo de estado actual del mundo, y las reglas son como las metas y creencias de un agente. Las reglas de condición-acción están imbuídas es un ciclo de observación-pensamiento-decisión-acción y son ejecutado apareando las condiciones de las reglas con los hechos en la memoria de trabajo y generando las acciones de esas reglas como candidatas a ser ejecutadas.

## Resumen del Capítulo 7: El Piojo y el Explorador de Marte

Esta forma de ejecutar es denominada encadenamiento hacia adelante y es similar al razonamiento hacia adelante. Si más de una acción candidata es generada, entonces se usa un procedimiento conocido como resolución de conflicto, para decidir entre las candidatas. La acción elegida es entonces ejecutada con lo que cambia el estado de la memoria de trabajo, simulando la forma en que las acciones del agente cambian el estado del mundo. Desde un punto de vista lógico, existen tres tipos de reglas de condición-acción: reglas reactivas, que son como asociaciones instintivas estímulo-respuesta; reglas de reducción de metas, que reducen metas a submetas usando encadenamiento hacia adelante; y reglas de razonamiento hacia adelante, que realizan el auténtico razonamiento homónimo.

## Resumen del Capítulo 8: Metas de mantenimiento como las fuerzas conductoras de la vida

El modelo de agente que se presenta en este libro combina las funcionalidades de la lógica y los sistemas de producción en un mismo marco lógico. El marco toma de los sistemas de producción el ciclo de observación-pensamiento-decisión-acción, pero reemplaza a las reglas de condición-acción con metas y creencias que tiene la forma lógica de los condicionales. El marco sustituye a las reglas reactivas con metas de mantenimiento que se usan para razonar hacia adelante, a las reglas de reducción de metas con creencias usadas para razonar hacia atrás, y a las reglas de razonar hacia adelante también con creencias que se usan para razonar hacia adelante. En el modelo lógico de un agente, el ciclo del agente responde a las observaciones que provienen del ambiente razonando hacia adelante con creencias, hasta que deriva una conclusión que se pueda aparear con una de las condiciones de las metas de mantenimiento.

## Resumen del Capítulo 8: Metas de mantenimiento como las fuerzas conductoras de la vida

Razona entonces hacia atrás, para chequear las otras condiciones de esa meta de mantenimiento. Si puede mostrar que todas las condiciones de esa meta de mantenimiento se cumplen, razona hacia adelante en un paso, derivando la conclusión de la meta de mantenimiento convertida en un meta de alcance o meta por lograr. A partir de esta meta, razona hacia atrás usando sus creencias para reducir la meta de alcance a un plan de acciones candidatas. Decide entonces entre las diferentes candidatas alternativas y comienza a ejecutar el plan. Si es necesario, interrumpe la ejecución del plan, para procesar otras observaciones, entretejiendo este plan con otros planes y acciones.

## Resumen del Capítulo 9: El significado de la vida

El marco lógico de los capítulos anteriores sugiere que la vida de un agente está controlada por los cambios que ocurren en el mundo, por sus propias metas y creencias, y por las decisiones que el agente toma entre las diversas maneras de alcanzar sus metas. La combinación de sus creencias y sus metas de más alto nivel generan una jerarquía de metas y submetas. Sin embargo, en procura de eficiencia, esta jerarquía puede ser colapsada en una colección de asociaciones estímulo-respuesta más directas, cuyas metas originales no son más aparentes, pero se mantienen implícitas y emergentes. En este capítulo, contrastamos la aparentemente absurda vida de un insecto imaginario, artificial, con la más significativa vida de un agente inteligente, en el cuál se combinan asociaciones de estímulo-respuesta con conciencia de las metas más elevadas.

## Resumen del Capítulo 10: Abducción

Una de las funciones principales de las creencias de un agente es representar las relaciones causales entre sus experiencias. El agente usa estas representaciones causales tanto proactivamente para generar planes para alcanzar sus metas, como preactivamente para derivar las consecuencias de acciones candidatas para ayudarse a escoger entre acciones alternativas. No obstante, el agente también puede las mismas creencias causales abductivamente, para generar hipótesis que expliquen sus observaciones, y para derivar las consecuencias de hipótesis candidatas al escoger entre explicaciones alternativas. Este proceso de generar y escoger hipótesis para explicar observaciones se denomina abducción. Como ocurre con el razonamiento por omisión de la negación por falla, la abducción es defectible en el sentido de que el arribo de nueva información puede causar que conclusiones derivadas previamente sean anuladas.

# Resumen del Capítulo 11: El dilema del Prisionero

El problema de decidir entre explicaciones abductivas de una observación es similar al problema de escoger entre cursos alternativos de acción, que tiene una excelente expresión en el Dilema del Prisionero. En este capítulo vemos como un agente puede usar una combinación de la lógica computacional y la teoría de decisiones para escoger entre alternativas. De acuerdo a las Teoría de Decisiones, el agente debe escoger la alternativa que le ofrece el mejor resultado esperado. El resultado esperado de una acción se calcula combinando apropiadamente juicios sobre la utilidad (o deseabilidad) de las consecuencias de una acción con juicios sobre la probabilidad (o verosimilitud) de que esas consecuencias de hecho ocurran.

# Resumen del Capítulo 11: El dilema del Prisionero

La teoría de decisiones es una teoría normativa, que demanda conocimiento detallado de las utilidades y las probabilidades, pero desprecia las motivaciones de las acciones del agente. En la práctica, los agentes típicamente optan por metas y creencias heurísticas (reglas de la experiencia), que aproximan las normas de la teoría de decisiones. Pero esas heurísticas algunas veces fallan. Cuando es importante tomar decisiones de mejor calidad, es mejor usar el más general marco del ciclo del agente, para analizar las motivaciones de las acciones y garantizar que todo un rango completo de alternativas es correctamente evaluado.



## Resumen del Capítulo 12: Las motivaciones importan

La Teoría de Decisiones conduce a teorías de moralidad consecuencialistas, que juzgan el estatus moral de las acciones simplemente como función de sus consecuencias. Pero en los estudios de psicología y en las leyes, las personas juzgan las acciones tanto en términos de sus consecuencias como en términos de sus motivaciones. Aquí mostramos como la lógica Computacional puede modelar esos juicios morales usando restricciones para impedir acciones que están condenadas a ser moral o legalmente inaceptables.

## Resumen del Capítulo 13: El mundo cambiante

La vida de un agente es una lucha constante para mantener una relación armoniosa con el siempre cambiante mundo. El agente asimila sus observaciones del estado cambiante del mundo y a cambio realiza acciones para cambiarlo. El mundo, por su parte, tiene vida propia. Existe en el presente, destruye su pasado y oculta su futuro. Para ayudarse a sobrevivir y prosperar en un ambiente así de cambiante, un agente inteligente usa creencias acerca de causas y efectos representadas en su lenguaje mental. En este capítulo investigamos en mayor detalle la representación lógica de esas creencias causales y la relación semántica entre esa representación lógica y el cambiante mundo.

## Resumen del Capítulo 14: Lógica y objetos

Mientras en la Psicología Cognitiva, los sistemas de producción son los competidores principales de la lógica, en Computación su principal competidor es la Orientación a los Objetos. En la forma orientada-a-objetos de ver el mundo, este consiste de objetos que interactúan enviando y recibiendo mensajes. Los objetos responden a los mensajes usando métodos encapsulados, invisibles para otros objetos y heredados de las clases más generales. La lógica computacional es compatible con la orientación por objetos si los objetos son vistos como agentes, los métodos son vistos como metas y creencias y los mensajes son vistos como provenientes de un agente que suministra información o que solicita ayuda de otros. Así considerados, la contribución principal de la orientación a los objetos es doble: Destacar el valor de organizar el conocimiento (las metas y las creencias) en módulos relativamente auto-contenidos y el valor de organizar ese conocimiento en jerarquías abstractas.

# Resumen del Capítulo 15: Bicondicionales

En este capítulo exploramos la perspectiva de las creencias condicionales como bicondicionales disfrazadas. Por ejemplo, dadas solamente las dos condiciones alternativas siguientes para hacer que un objeto parezca rojo:

Un objeto luce rojo si el objeto es rojo. Un objeto luce rojo si está iluminado por una luz roja.

las dos condicionales puede asumirse que representan a la bicondicional:

Un objeto luce rojo si y solo si el objeto es rojo o el objeto está iluminado por una luz roja.

Tanto la negación por falla como la abducción se pueden entender como formas de razonamiento que usan esas bicondicionales como si fuese equivalencias, reemplazando fórmulas atómicas que se aparean con la conclusión con la disyunción de condiciones (conectadas por el o) que implican la conclusión.

# Resumen del Capítulo 16: Lógica Computacional y la Tarea de Selección

En este capítulo volvemos al problema de explicar algunos de los resultados los experimentos psicológicos acerca de razonar con condicionales. Investigamos las diferentes maneras en la que la Lógica Computacional explica esos resultados y que dependen de que una condicional sea interpretada como meta o como creencia. Si se le interpreta como una creencia, con frecuencia es natural interpretar las condicionales como si especificaran las únicas condiciones bajas las cuáles se cumple la conclusión.

# Resumen del Capítulo 16: Lógica Computacional y la Tarea de Selección

Esto explica uno de los dos tipos principales de error que la gente comete cuando razonan con condicionales, vistos desde la perspectiva de la lógica clásica. El otro error principal es que la gente suele equivocarse con la negación. Este error es explicable, en parte, por el hecho de que las observaciones del agente normalmente se representan con oraciones atómicas positivas, y que las conclusiones negativas se tienen que obtener a partir de observaciones positivas. En muchos casos, esta derivación es más fácil de hacer con metas condicionales que con creencias condicionales

# Resumen del Capítulo 17: Metalógica

En este capítulo exploramos el cómo se puede usar la meta-lógica para simular el razonamiento de otros agentes y para resolver problemas que no pueden ser resueltos en el lenguaje objeto puro. Ilustramos esta posibilidad con una variante del acertijo de los sabios y con el teorema de Gödel que establece que hay oraciones ciertas pero que no puede ser probadas en la aritmética.

Este capítulo de conclusiones se abstrae de los detalles y adopta punto de vista más amplio respecto al objetivo principal de este libro, que es mostrar que la Lógica Computacional puede reconciliar paradigmas conflictivos para explicar y guiar la conducta humana. También sugiere que la Lógica Computacional puede ayudar a resolver conflictos en otras áreas.



# Resumen del Capítulo A1: La sintaxis de la forma lógica

Este capítulo adicional, más formal que los anteriores, presenta una formulación más precisa de la Lógica Computacional como una lógica de oraciones que tienen la forma condicional: *si condiciones entonces conclusión* o, la forma equivalente, *conclusión si condiciones*. En su forma más simple, la *conclusión* de una condicional es una *expresión atómica*, que consiste de un *predicado* y una serie de *argumentos*. Las *condiciones* son una conjunción (conectada por la palabra *y*) de expresiones atómicas o las *negaciones* de expresiones atómicas.

# Resumen del Capítulo A1: La sintaxis de la forma lógica

En este capítulo, comparo la forma condicional de la lógica con la lógica clásica estandar, Argumento que la lógica clásica es a la lógica condicional, como el lenguaje natural es al lenguaje del pensamiento. En ambos casos, hay dos clases de razonamiento, desplegado en dos etapas. La primer etapa traduce las oraciones que no tienen estructura y posiblemente sean difíciles de entender en oraciones más simples con (mejor) estructura. En la segunda etapa se derivan las consecuencias de esas oraciones más simples. La lógica de las formas condicionales es la lógica de esas oraciones mejor estructuradas y más simples.

## Resumen del Capítulo A2: La Verdad

Los condicionales en lógica computacional representan las metas y creencias de un agente en su lenguaje privado para pensar. También representan los significados de sus comunicaciones, públicas vale decir, con otros agentes, y por esta razón se puede decir que representan la *semántica* de las oraciones en lenguaje natural. Sin embargo, las oraciones en forma lógica tienen a su vez una *semántica* definida en términos de su relación con los estados del mundo.

# Resumen del Capítulo A2: La Verdad

Este capítulo adicional comienza la discusión sobre esta semántica (la de las formas condicionales respecto al estado del mundo) y sobre la relación entre la verdad en todos los modelos y la verdad en los modelos mínimos. Se argumenta, a partir del ejemplo de la aritmética, que la verdad en los modelos mínimos es más fundamental que la verdad en todos los modelos.

# Resumen del Capítulo A3: Razonamiento hacia adelante y hacia atrás

Este capítulo define con más precisión las reglas de inferencia para razonar hacia adelante y hacia atrás, y demuestra cómo se les puede entender en términos semánticos, mostrando como la verdad de un conjunto de oraciones implica la verdad de otro conjunto. Este punto de vista semántico se aplica tanto al uso de esas reglas de inferencia para determinar la verdad en todos los modelos, como a su uso para generar y determinar la verdad en los modelos mínimos.

Este capítulo demuestra como la semántica de la negación por falla puede entenderse en términos de la semántica de los modelos mínimos del capítulo A2.

## Capítulo A5: La Regla de Inferencia Llamada Resolución

En este capítulo, se considera tanto al razonamiento hacia adelante como al razonamiento hacia atrás como casos especiales de la regla de inferencia conocida como Resolución y se explica que resolución es el mecanismo subyacente en el razonamiento en los grafos de conexión.

A Resolución se le presentó originalmente como una regla de inferencia orientada a la máquina, mientras que los razonamientos hacia adelante y hacia atrás son maneras dirigidas a *humanos* de entender el pensamiento humano. Esta combinación de las orientaciones hacia la máquina y hacia el humano se refleja en el hecho de que la mente humana puede considerarse una computadora cuyo software se escribe en la forma condicional de la lógica y cuyo hardware es una forma conexionista de resolución.

# Resumen del Capítulo A6: La lógica de la Programación Lógica Abductiva

Este capítulo ofrece la mayor parte del soporte necesario para entender la combinación de razonamiento hacia adelante, hacia atrás y la negación por falla, que son las reglas de inferencias básicas de la Logica Computacional que se usa en este libro.

El procedimiento de prueba que se presenta en este capítulo se puede entender en términos semánticos, como el proceso de generar un modelo mínimo en el cual todas las metas y creencias de un agente son verdaderas. Sin embargo, también se le puede entender en términos de argumentación, como la generación de un argumento a favor de una declaración, que le da apoyo directo a esa declaración, pero que también refuta o vence, con contra argumentos, los argumentos que atacan a esa declaración.