

Capítulo I

Conocimiento

I.1. Razonamiento Conceptual

El conocimiento se constituye por los conceptos generales y particulares en la mente de agentes plenamente racionales. Se puede expresar de forma pública a través del lenguaje y puede ser objeto de la introspección; por lo mismo tiene un carácter simbólico y declarativo, y se distingue de la información que habilita las habilidades perceptuales y motoras, que se embeben de forma sub-simbólica en las estructuras respectivas.

Se caracteriza asimismo porque el agente concibe al mundo como constituido por entes individuales, tanto concretos como abstractos, con sus propiedades y relaciones, que se presentan a la mente como un producto acabado de la interpretación perceptual y/o del pensamiento. El conjunto de estas entidades constituye el *universo*, *dominio* u *ontología* de individuos acerca de los que se tiene conocimiento.

Esta conceptualización se puede elaborar con otro tipo de entidades cuya extensión espacial y temporal es difusa, como los fluidos o los grumos, en los que no se distinguen claramente las partes del todo, o como los eventos, las acciones y los procesos, pero la noción de entidad individual, que tiene un conjunto de propiedades particulares y establece relaciones con otras entidades, incluyendo a sí misma, es la intuición básica.

Las entidades individuales se presentan directamente a la consciencia de los seres humanos, y posiblemente de otros animales con un sistema nervioso suficientemente desarrollado, y pueden ser objeto de la atención –y señalarse con gestos ostensivos– así como de la manipulación mecánica y de la afectividad. De forma recíproca, los individuos o especies que no dividen al mundo en entidades individuales se enfrentan al entorno de forma holística e indiferenciada. Es posible también que esta distinción sea gradual y el grado de individuación sea incremental, y que en el plano ontológico se desarrolle y madure a lo largo del desarrollo mental. En todo caso, la capacidad de *individuación* depende de la dotación perceptual innata y del pensamiento imaginativo, y se puede modelar hasta cierta instancia con programas de visión computacional e interpretación del lenguaje natural, por ejemplo, pero su comprensión cabal es un misterio de la mente.

Aunada de manera muy cercana a la intuición básica de la entidad individual es la noción de *clase*. Ésta consiste en la partición del dominio en regiones mutuamente excluyentes, donde cada región contiene a individuos suficientemente similares desde la perspectiva conceptual que se adopte en el análisis. Estas sub-particiones son las *subclases* del dominio. Las relaciones entre subclases y clases es de inclusión, y la de individuos y clases de membresía. Por ejemplo, el universo de todos los entes se puede partir primero entre los abstractos y los concretos; y los concretos entre los vivos y los inanimados. En esta jerarquía, la clase de los seres vivos se incluye en la clase de los objetos concretos y esta última en la de todos los entes; y un ser vivo particular es miembro de las clases de los seres vivos, de los concretos y de los entes. Por supuesto, la ontología puede variar con el tiempo y en relación a los mundos posibles, y estas consideraciones se tienen también que tomar en cuenta.

La formulación de la ontología y su clasificación depende del dominio de conocimiento que se aborde, y una premisa básica es que el dominio sea completo, es decir, que incluya a todos los individuos materiales o abstractos del espacio del problema bajo estudio; de otra forma los modelos del conocimiento serían incompletos.

Los individuos tienen propiedades y relaciones. Por ejemplo, los objetos materiales tienen una extensión en el espacio y un ente puede estar adelante o atrás de otro. La extensión es una propiedad y estar enfrente o atrás son relaciones. Las propiedades pueden ser boo-

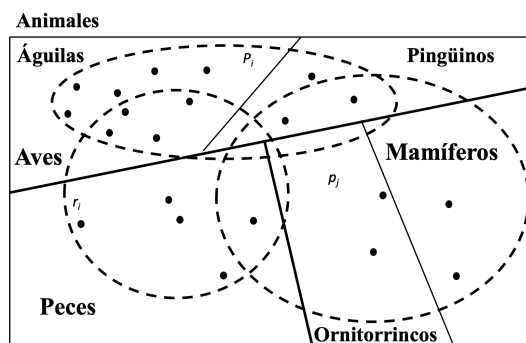


Figura 1.1: Individuos, clases, propiedades y relaciones

leanas, como si un ave vuela o no vuela, o pueden tener otro tipo de valores, por ejemplo, si su color es blanco, donde “color” es la propiedad y “blanco” el valor. A diferencia de los individuos que de manera fundamental pertenecen a una clase y las clases que la contienen, las propiedades y las relaciones se puede adscribir a individuos de clases diferentes de manera muy arbitraria. Por ejemplo, los pingüinos y las águilas pueden compartir la propiedad de tener pecho blanco y cuerpo negro, aunque sean de diferentes clases.

La Figura 1.1 ilustra la ontología de los animales donde cada punto representa a un animal particular y las regiones demarcadas por líneas continuas a las clases y, en particular, el rectángulo exterior representa al dominio o universo. Las líneas más gruesas dividen al espacio en tres subclases principales que se indican con las letras de mayor tamaño: las aves, los peces y los mamíferos; y las clases en el siguiente nivel se demarcan con líneas más delgadas y letras más pequeñas: las aves se dividen águilas y pingüinos, y se distinguen los ornitorrincos de los demás mamíferos. Los óvalos con líneas punteadas representan a las propiedades p_i , p_j y la relación r_i e indican que los puntos que contienen representan a individuos que tienen las propiedades o entran en las relaciones correspondientes. Los óvalos las trascienden las fronteras de las regiones que representan a las clases arbitrariamente. El diagrama ilustra que los individuos son la intuición básica, que precede a las clases, y que las propiedades y las relaciones son más contingentes.

1.2. Representación del Conocimiento

El diagrama es una representación del conocimiento. Ésta se presenta de manera pública para su interpretación por parte de los seres humanos, y se distingue de la ontología propiamente, que se constituye por los objetos en el mundo. La representación diagramática se puede expresar también como una taxonomía estricta y representar computacionalmente por medio de jerarquías análogas a las que se utilizan en el algoritmo Minimax y programas similares, como se ilustra en la Figura 1.2. Los círculos representan a las clases y las flechas la relación de contención; los nombres de las clases se incluyen en negrillas en el círculo correspondiente, y las etiquetas asociadas a los círculos indican las propiedades que tienen todos los miembros de la clase. El símbolo $\Rightarrow$ en las etiquetas de forma $\alpha \Rightarrow \beta$ indica que la clase representada por el círculo está en la relación α con la clase β . La lectura es genérica y no se especifica quién tiene la propiedad o entra en la relación. Por ejemplo, en la Figura 1.2 se expresa que las aves vuelan, que los pingüinos nadan, que los ornitorrincos ponen huevos y que las águilas comen peces.

Asimismo los individuos particulares se representan por un rectángulo que incluye sus propiedades y relaciones, y las flechas que relacionan círculos con cuadrados representa la relación de membresía. En este caso la etiqueta $\alpha \Rightarrow \beta$ indica que el individuo representado por el cuadrado tiene la propiedad α cuyo valor es β , o que está en la relación α con el individuo β . En este caso la lectura es específica ya que las propiedades y relaciones son de y se establecen entre individuos concretos. El diagrama indica que Pedro es grande y que Arturo es listo y es amigo o está en la relación de amistad con Pedro.

Las relaciones de inclusión y membresía representan a la relación de herencia o *modus ponens* y todos los individuos tienen además de sus propiedades y relaciones particulares las de su clase y sus súper clases. Por ejemplo, el diagrama representa implícitamente que Pedro come peces y vuela; y que Arturo nada y vuela. Las propiedades y relaciones de una clase se pueden ver también como valores asumidos o *defaults* que todos los miembros de la clase, ya sea por estipulación directa o por la relación de herencia, tienen.

La economía es una propiedad muy importante de las representaciones, que en general deben tener un buen compromiso entre lo que se expresa explícita versus implícitamente.

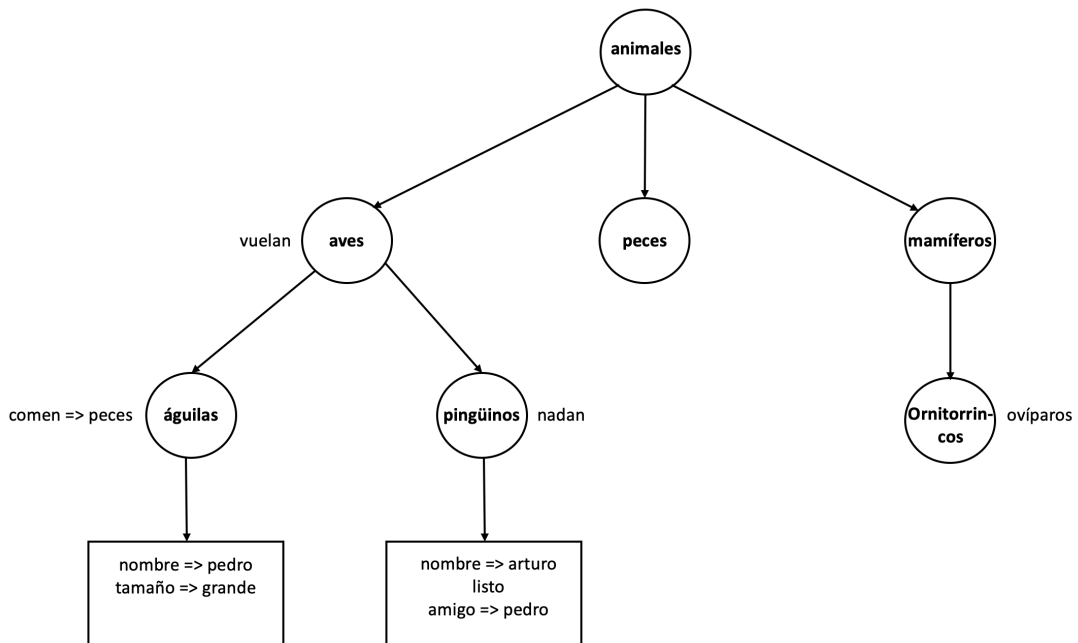


Figura 1.2: Taxonomía de los animales

Éste es también un compromiso entre la cantidad de memoria que se requiere para almacenar a la información y el esfuerzo o costo del proceso de cómputo para retribuirla. En el caso límite, si toda la información se expresara explícitamente se retribuiría por inspección directa; sin embargo, el conocimiento conceptual es un objeto muy grande y su almacenamiento explícito no es en general una opción viable. La partición del dominio en clases se motiva por un criterio de economía, para tener un compromiso más satisfactorio entre la memoria y el proceso.

Sin embargo, la estructura de la representación tiene efectos en la interpretación que pueden estar en conflicto con lo que se expresa explícitamente. Por ejemplo, la jerarquía conceptual en la Figura 1.2, bajo las convenciones de interpretación expuestas, expresa que los pingüinos, y en particular Arturo, vuelan, lo cual es falso. El diagrama representa una fase de la adquisición del conocimiento en la que ya se sabe que las aves vuelan y que los pingüinos son aves que nadan, pero todavía no se ha aprendido que no vuelan. Para expresar este hecho

se requiere poder expresar la negación; pero este incremento de la capacidad expresiva trae también conflictos interpretativos cuya resolución no es trivial.

1.3. Negación y Razonamiento No-Monotónico

Se conoce como *extensión* de la representación o de la *Base de Conocimiento* a la totalidad del conocimiento expresado explícita e implícitamente. En el caso de la taxonomía estricta el conocimiento implícito es todo lo que se sigue de la relación de herencia. La forma más básica de expresar la negación es asumir que lo que no se expresa es falso. Por ejemplo, si se pregunta si los peces vuelan en relación a la Figura 1.2, la respuesta será *no*, ya que la propiedad de volar no se estipula para los peces ni para los animales, su única súper clase, y en este caso es correcta.

A esta hipótesis de interpretación se le conoce como *Hipótesis del Mundo Cerrado* o *Closed World Assumption*. De manera más general ésta estipula que si algo no se sigue de la información expresada explícitamente y de los esquemas de inferencia del sistema de representación se considera que es falso. También se conoce como *negación bajo falla* en sistemas de prueba de teoremas y, en particular, en el lenguaje de programación Prolog. También se refiere en algunos contextos como *negación débil*.

La hipótesis del mundo cerrado es muy útil ya que normalmente se expresan las propiedades y relaciones que se tienen, y sólo se expresan las que no se tienen cuando es necesario marcarlas –además de que normalmente las propiedades y relaciones que un individuo no tiene, entre la totalidad posible, es mucho mayor que las que si tiene. Sin embargo, es posible equivocarse; por ejemplo, si se pregunta si los peces comen animales, la respuesta será igualmente negativa, pero en este caso falsa. Consecuentemente, la hipótesis del mundo cerrado sólo se puede aplicar cuando se asume que la información expresada en la representación es completa; es decir que no hay hechos relevantes del dominio que no se hayan incluido. Sólo en este caso al inferencia se puede considerar como válida o correcta, y de ahí el nombre de la hipótesis.

La negación es un gran salto en la representación del conocimiento, pero también trae consigo conflictos muy severos. En la Figura 1.3 se introduce la negación y se incluyen algunas

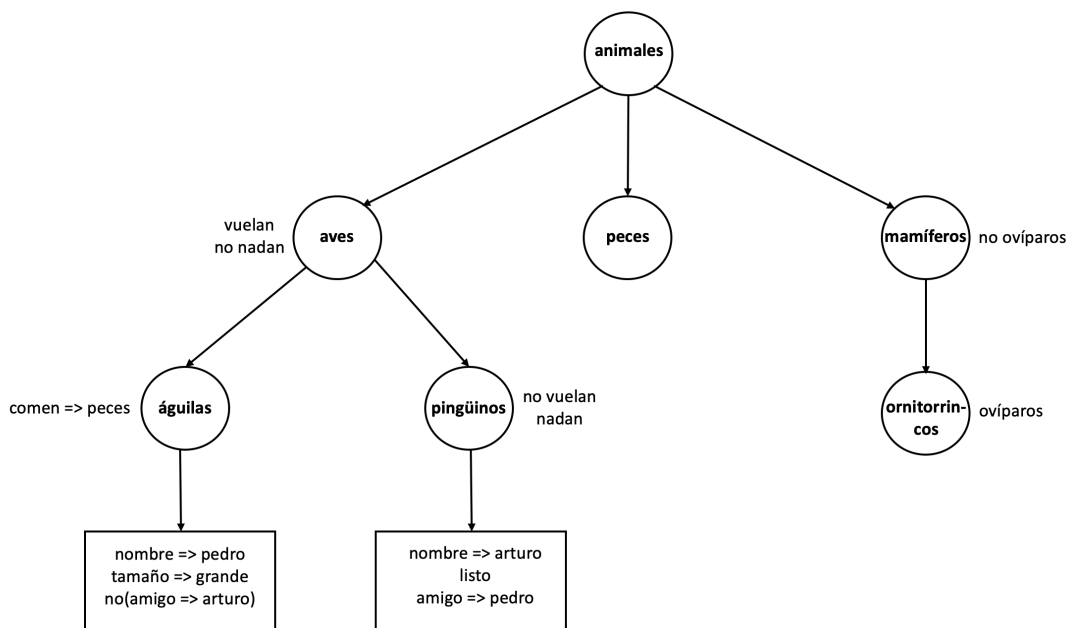


Figura 1.3: Razonamiento no-monotónico

proposiciones negativas relevantes, en particular que Pedro no es amigo de Arturo, lo que origina un pequeño drama. Se incluye además que los pingüinos no vuelan, que las aves no nadan y que los mamíferos no ponen huevos. La negación explícita, en oposición a la negación bajo falla, expresa la certeza que la proposición que se niega es falsa. Para distinguirla de la negación débil a veces se refiere como *negación fuerte*. Esta es la negación del lenguaje cotidiano y de los lenguajes lógicos.

La negación explícita permite razonar acerca de dominios de conocimientos abiertos, y garantizar que las inferencias sean válidas aún si la representación es incompleta. En este tipo de sistemas la respuesta a una pregunta deberá estar en la extensión de la base de conocimientos, tanto si la proposición es positiva como negativa, y en caso contrario el sistema deberá responder *no sé*. Si se pregunta, por ejemplo, si las águilas nadan en relación a la Figura 1.3 la respuesta será *no*, pero si se pregunta si los peces vuelan en relación a la misma figura la respuesta deberá ser *no sé*.

Es posible también utilizar la hipótesis del mundo cerrado incluso en caso de que no se sepa, y responder *no* en vez de *no sé*. Esta conducta es psicológicamente plausible y aunque se confunde la negación fuerte con la débil, las respuestas podrían ser correctas frecuentemente, ya que hay mucho más proposiciones negativas que positivas, pero también es posible equivocarse con consecuencias muy costosas. Por otra parte, hay una tendencia a la creatividad y la cortesía mexicana que se resiste a decir no sé y se prefiere responder que sí aunque no se sepa. Por ejemplo, ante la pregunta si los elefantes comen mariposas se respondería que sí, seguida de una discusión *ad nauseam*.

La riqueza expresiva que provee la negación tiene que enfrentar que las representaciones se pueden volver inconsistentes. Bajo las condiciones de interpretación estipuladas, la taxonomía en la Figura 1.3 expresa que los pingüinos –en particular Arturo– vuelan y no vuelan, que nadan y no nadan, y que los ornitorrincos ponen y no ponen huevos. Este es un problema muy serio para cualquier sistemas representacional porque se puede siempre concluir cualquier cosa aunque el razonamiento no sea válido; por ejemplo, la proposición que los pingüinos no vuelan se apoya en la expresión directa de este hecho, pero su negación se apoya en que los pingüinos son aves y volar es una propiedad que tienen todas las aves.

Estas inconsistencias surgen de la estructura de la representación y para abordarlas hay que introducir convenciones, heurísticas o preferencias de interpretación adicionales que permitan afrontar el problema, aunque sea sólo parcialmente. Una heurística muy útil es el llamado *Principio de Especificidad* de acuerdo con el cual en caso de conflictos de conocimiento se prefiere siempre a la proposición más específica. Los individuos son más específicos que las clases, y las clases más específicas corresponden a las regiones más anidadas de la ontología, o a los nodos más inferiores de la taxonomía.

Bajo este principio el diagrama en la Figura 1.3 expresa que los pingüinos, en particular Arturo, no vuelan pero nadan, y que los ornitorrincos ponen huevos. Las propiedades generales como que las aves vuelan y no nadan, y que los mamíferos no ponen huevos, se mantienen pero “se bloquean” para las excepciones, que corresponden a la información más específica. Los defaults y las excepciones pueden ser tanto positivos como negativos; por ejemplo *vuelan* es un default positivo y los pingüinos es la excepción negativa, pero *no nadan* es una default

negativo y los pingüinos son la excepción positiva. Con estas consideraciones la interpretación de la taxonomía se hace consistente. A este tipo de razonamiento se le conoce como *no monotónico* y es una aportación de la Inteligencia Artificial a la lógica [18].

El término *no monotónico* proviene de la negación de una propiedad general del conocimiento lógico y matemático; ésta consiste en que el valor de verdad de una proposición es permanente: si algo es verdadero es verdadero para siempre. Por ejemplo, cuando se prueba un teorema matemático, éste es verdadero necesariamente. Si posteriormente se muestra que el teorema es falso, es una falla de la prueba original, pero no de la proposición propiamente. Sin embargo, el conocimiento conceptual que tiene una base empírica no tiene esta propiedad y el valor de verdad de las proposiciones pueden cambiar porque se incrementa nuestro conocimiento, como cuando el niño aprende que los pingüinos son aves que no vuelan, o simplemente porque el mundo cambia todo el tiempo. Los sistemas de representación del conocimiento así como la memoria conceptual humana tienen que ser flexibles para acomodar estos cambios y reflejar al mundo lo más precisamente posible.

1.4. Preferencias, Restricciones y Justificaciones

Las contradicciones que dan lugar al razonamiento no monotónico dependen de la interacción de la negación con la estructura jerárquica de la taxonomía, pero pueden ocurrir por otras razones. En particular los defaults especificados hasta este punto son absolutos y se aplican necesariamente, pero es posible condicionarlos para que sólo se aplique en contextos específicos. Para este efecto se requiere aumentar la expresividad del sistema representacional con la implicación lógica, la cual se representa por las expresiones de forma $\Gamma \Rightarrow \Delta$, donde $\Rightarrow$ es el símbolo de implicación, Γ y Δ son conjuntos de proposiciones, el antecedente y el consecuente respectivamente. Esta nueva expresividad se ilustra en la Figura 1.4 donde el ejemplo se enriquece con las etiquetas $\{trabaja(y)\} \Rightarrow \{vive(y)\}$ y $\{nació(y)\} \Rightarrow \{vive(y)\}$ asociadas a las clase *animales*, y es una variables que se refiere a individuos particulares. Su interpretación es *si cualquier animal trabaja en y entonces vive en y* y *si cualquier animal nació en y entonces vive en y* respectivamente. En el ejemplo todas las proposiciones

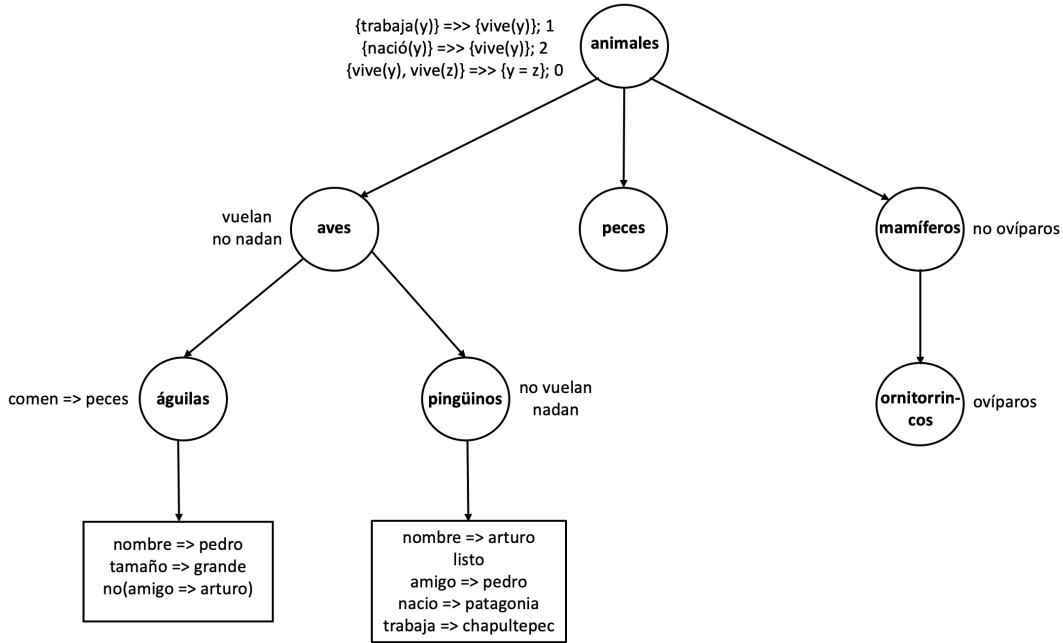


Figura 1.4: Preferencias y Justificaciones

son positivas, pero no hay restricciones y se pueden incluir proposiciones negativas tanto en las premisas y en el consecuente.

La figura 1.4 muestra un nuevo estado de la representación en el que el agente ha adquirido o aprendido esta nueva información. En la figura se ilustra también que Arturo, el pingüino, nació en Patagonia y que trabaja en Chapultepec. Estos hechos se aprenden a través del lenguaje y posiblemente otras modalidades como la visión. La propiedad condicional se hereda a todas las clases abajo de *animales* y consecuentemente a las aves y a los pingüinos, y la extensión de la base de conocimiento incluye ahora dónde vive Arturo, pero no se aplica a Pedro, ya que para éste no se especifica dónde nació ni dónde trabaja.

Sin embargo, el aumento de la expresividad tiene también un costo y viene acompañado de un nuevo tipo de incoherencias y/o contradicciones. Se sigue en particular que Arturo vive en Chapultepec y en la Patagonia. Esto es directamente inocuo, pero si se establece la proposición de sentido común que un individuo no puede vivir en más de un lugar al mismo tiempo, las dos proposiciones no se pueden sostener simultáneamente. Para expresar esta

última hay que incrementar nuevamente la expresividad del sistema con la igualdad entre individuos, la cual se expresa con el símbolo $=$. La expresión $\{vive(y), vive(z)\} \Rightarrow \{y = z\}$ asociada a animales establece este default genérico; es decir, si cualquier animal vive en el lugar denotado por y así como por z , el valor estas dos variables tiene que ser el mismo. A este tipo de preferencias con peso cero les llamamos *restricciones*. De acuerdo con esta restricción una de las preposiciones *Arturo vive en Chapultepec* y *Arturo vive en la Patagonia* se tiene que descartar.

El principio de especificidad no es útil para hacer esta distinción ya que los tres defaults condicionales están al mismo nivel o son igualmente específicos, y es necesario acudir a un criterio adicional para resolver este problema. Para este efecto se introduce el factor de preferencia, peso o prioridad, el cual se indica con el número a la derecha de los defaults condicionales. Este se estipula aquí pero se puede aprender por experiencia. Se adopta la convención de que mientras más bajo es más prioritario. Se asume asimismo que todas las proposiciones tienen una prioridad, y las propiedades y relaciones tienen prioridad cero. Nos referimos aquí a toda proposición con un peso mayor que cero como *preferencia*. Los conflictos de conocimiento que no se pueden resolver por el principio de especificidad se resuelven por su valor de preferencia. En el ejemplo se concluye que Arturo vive en Chapultepec porque se prefiere sobre a que vive en la Patagonia. La intuición es que es más plausible que alguien viva donde trabaja que donde nació, aunque esto es contingente y las preferencias pueden cambiar con la experiencia.

Las implicaciones se pueden utilizar también de forma inversa para producir justificaciones o explicaciones de los hechos que se observan en el mundo. Supongamos, por ejemplo, que el agente, que tiene la representación en la Figura 1.4 en su memoria, aprende, ya sea porque lo vea o se lo digan, que Arturo vive en Chapultepec, y se le pregunta y/o quisiera saber o justificar este hecho. Para este efecto es posible utilizar las implicaciones en reversa, y hacer las hipótesis más plausible de las causas de los hechos observados. La explicación en este caso es porque ahí trabaja, ya esta proposición es más prioritaria que porque ahí nació. Las inferencias en reversa pueden estar encadenadas, y la premisa de una implicación puede ser la consecuencia de otra, y se puede generar un árbol con varias trayectorias “hacia atrás”,

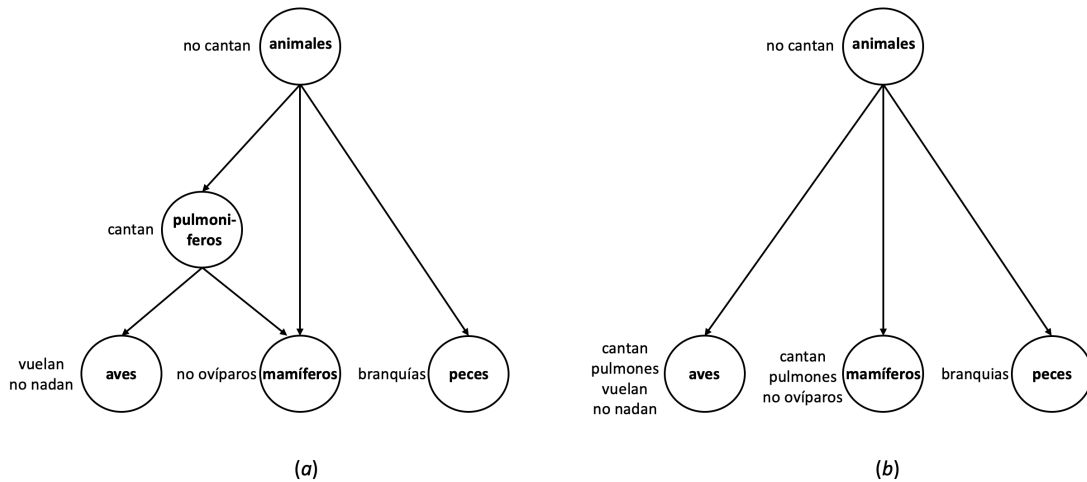


Figura 1.5: Látices y herencia múltiple

cada una de las cuales representa una justificación posible. La prioridad de cada hipótesis será función los pesos de las implicaciones utilizadas en la misma. Esta inferencia en reversa, de las consecuencias a los antecedentes, corresponde al diagnóstico o la abducción del ciclo de inferencia de la vida cotidiana.

1.5. Látices y Herencia Multiple

La expresividad del sistema representacional se incrementa mediante la inclusión operadores como la negación, la implicación y la igualdad, pero es también posible aumentarla a nivel estructural. Por ejemplo, la jerarquía estricta de la taxonomía se puede relajar admitiendo retículas o *látices* en las que una clase puede tener más de una madre. Por ejemplo, en la Figura 1.5 (a) se extiende la parte superior de la ontología de animales con la clase *pulmoníferos*, de acuerdo a la intuición de que los animales que tienen pulmones constituyen una clase, de la misma forma que los que tienen mamas.

Sin embargo, esta extensión tiene el costo de que la relación de herencia se puede dar por varias trayectorias y se puede heredar una proposición por una ruta y su negación por otra. En la Figura 1.5 (a) se expresa que los animales no cantan pero los pulmoníferos son un

excepción a este default negativo. Asimismo, la clase de animales domina directamente a los pulmoníferos y a los mamíferos, y la de pulmoníferos a las aves y los mamíferos. Se sigue que las aves cantan, pero también que los mamíferos cantan –por ser pulmoníferos– y no cantan –por ser animales.

Este conflicto se puede abordar con el principio de especificidad pero en una forma más elaborada y el mecanismo es mucho más complejo. Por ejemplo, se podría proponer la heurística que la propiedad que se hereda proviene de la trayectoria más larga a partir del nodo más específico que es común a todas las trayectorias que llegan a la clase donde surge el conflicto. En el caso del ejemplo se concluiría que los mamíferos cantan porque esta propiedad se hereda por una trayectoria de longitud dos, y se prefiere a que no cantan, que se hereda por una de longitud uno. En esta forma elaborada del principio de especificidad las trayectorias más largas se consideran más específicas.

Sin embargo, la contradicción se disuelve si se considera que tener pulmones es una propiedad y no una clase, que se adscribe a las aves y a los mamíferos, en conjunto con la propiedad de cantar, como se muestra en la Figura 1.5 (b). Con esta consideración la estructura vuelve a ser una taxonomía estricta, no hay herencia múltiple y los defaults se resuelven por el principio de especificidad. Este ejemplo refuerza la intuición que los individuos y las clases son conceptualizaciones más primitivas que las propiedades y las relaciones, y que adoptar la perspectiva alterna –que las clases se determinan o definen en relación a un conjunto de *propiedades esenciales*– complica la estructura innecesariamente y conlleva un alto costo inferencial.

La combinación de la jerarquía aunada a las etiquetas textuales constituye un formato representacional con una estructura simple pero que permite representar dominios de conocimiento ricos, y presenta un buen compromiso entre expresividad y requerimientos de memoria y esfuerzo de cómputo. Si embargo, hay situaciones en que se requiere aumentar la expresividad estructural y emplear látices o incluso otras estructuras de mayor complejidad.

La taxonomía corresponde a la partición del dominio en regiones mutuamente excluyentes, mientras que la látice admite que las regiones estén traslapadas. Esta situación se puede conceptualizar como un conjunto de particiones paralelas –mutuamente excluyentes– que se

proyectan en un sólo plano. Cada plano contiene a una perspectiva perceptual particular con un dominio y una partición mutuamente excluyente, y la proyección contiene la superposición de los dominios y particiones, y representa a la integración de las diferentes perspectivas.

Por ejemplo, en la interpretación de los planos de una casa, aunque las entidades básicas –los cuartos y pasillos con sus paredes, puertas, tuberías y conductos eléctricos, que se representan por áreas y líneas– son las mismas o muy similares para el albañil, el electricista, el plomero, el diseñador de interiores, etc., cada quien tiene una perspectiva conceptual que atiende a los objetivos y prácticas de su oficio. Cada perspectiva conceptualiza un dominio con los individuos relevantes, y lo parte en regiones mutuamente excluyentes, aunque puede haber una intersección significativa entre los diversos dominios. Pero el arquitecto debe proyectar las diversas perspectivas a un sólo plano y resolver los conflictos que sólo se pueden apreciar en la visión global, pero que se reflejarían en la casa, e incomodarían profundamente a la dueña.

1.6. Estados, Procesos, Acciones y Eventos

El conocimiento conceptual incluye también estados, procesos, acciones y eventos cuya representación es muy compleja y elusiva, y se tiene que considerar primero si éstos se pueden conceptualizar como objetos individuales de naturaleza abstracta, como los números, pero que tienen permanencia espacial y temporal. Intuitivamente se pueden referir mediante el lenguaje, se pueden registrar y recordar en la memoria, y pueden ser objeto del razonamiento, de la misma manera del resto de los objetos y clases, y consecuentemente se pueden admitir en la ontología. Para este efecto se puede postular, por ejemplo, que la ontología consiste de la clase universal de todos los entes, que se parte en las clases disyuntivas de entes concretos y entes abstractos; la primera se parte a su vez en inanimados y animados o vivos, y la clase de los animales es una de las particiones de los seres vivos; por su parte la clase de entes abstractos se puede partir en los objetos puros del pensamiento, como los objetos matemáticos, y aquellos que tienen una extensión espacial y temporal, como son los estados, procesos y eventos naturales, así como las acciones reactivas e intencionales realizadas por los organismos más

básicos y los seres vivos con un sistema nervioso suficientemente desarrollado respectivamente. Por supuesto, esta partición se puede hacer desde diversas perspectivas y se pueden producir diversas ontologías que corresponderán a la visión del mundo del analista.

Por ejemplo, extendiendo nuestro ejemplo, supongamos que se tiene acceso a la siguiente escena, ya sea porque la veamos directamente o porque no las cuentan: Arturo está sentado en un cubo de hielo tomándose un café mientras Pedro vuela dando círculos en el cielo; de pronto Pedro ve Arturo y se lanza sobre él con no muy buenas intenciones. Arturo se da cuenta, toma una piedra y se la avienta con buena puntería. El impacto le fractura el ala a Pedro quien cae al suelo. Pedro se convierte entonces en una águila que no vuela. Pedro se enoja en extremo y empieza a perseguir a Arturo sobre el hielo. No sabemos cuál es el final de esta historia. Para representar esta historia es necesario identificar a los estados e_i , los procesos p_j , las acciones a_k y los eventos naturales n_l como sigue:

- $e_1 \Rightarrow$ Arturo está sentado en un cubo de hielo;
- $p_1 \Rightarrow$ Arturo se está tomando un café;
- $p_2 \Rightarrow$ Pedro está volando en círculos en el cielo;
- $a_1 \Rightarrow$ Pedro se lanza sobre Arturo con malas intenciones;
- $a_2 \Rightarrow$ Arturo ve a Pedro;
- $a_3 \Rightarrow$ Arturo le lanza a Pedro una piedra con buena puntería;
- $n_1 \Rightarrow$ Piedra impacta el ala;
- $n_2 \Rightarrow$ Pedro se fractura el ala;
- $n_3 \Rightarrow$ Pedro cae al suelo;
- $e_2 \Rightarrow$ Pedro está muy enojado;
- $p_3 \Rightarrow$ Pedro persigue a Arturo sobre el hielo;

Convencionalmente y para fines de la exposición podemos establecer que los individuos de estas clases tienen como única propiedad su descripción. En el presente ejemplo éstas deben incluir las referencias apropiadas a los individuos con sus propiedades y relaciones referidos en la ontología en la Figura 1.4 y se requiere adicionalmente que se establezcan las clases e individuos de cubo de hielo, bebidas y café, objetos materiales inanimados y piedras, etc.

Estas entidades abstractas tienen también relaciones, cuya caracterización no es tampoco trivial. Su análisis detallado va más allá de los propósitos de este texto y aquí sólo consideramos las relaciones temporales y las causales; por ejemplo, que un estado o un proceso pueden preceder, traslaparse, coincidir o seguir temporalmente a otro; que una acción o un evento pueden ser causa de un estado o un proceso, etc. La escena del ejemplo tiene las siguientes relaciones:

- *traslapados*: $(e_1, p_1); (e_1, p_2); (p_1, p_2); (e_1, a_1); (p_1, a_1); (p_2, a_1); (a_1, a_2); (a_2, a_3);$
- *causa*: $(a_3, n_1); (n_1, n_2); (n_2, n_3); (n_3, e_2); (n_2, p_3);$

El presente análisis no pretende ser exhaustivo, y la interpretación de estas proposiciones se realiza en relación a un contexto de sentido común, que tampoco se hace explícito. El análisis depende también de la granularidad de las descripciones, y se tienen que realizar en relación a la perspectiva, los intereses, las creencias, etc., del narrador. Es necesario considerar adicionalmente si esta información se proporciona directamente en el proceso de análisis, o es producto de la relación del agente computacional con el entorno. Por ejemplo, si se induce en forma automática a partir de descripciones lingüísticas textuales, o a partir de percibir la escena visualmente.

Esta información permite en principio razonar acerca de las relaciones temporales y causales de la trama, y esta forma de razonamiento conceptual se puede asimismo retribuir durante el uso del lenguaje y el pensamiento. Las representaciones se almacenan en la memoria, y su uso constituye una forma de razonamiento conceptual, a la par que el razonamiento acerca de la ontología de los animales. Sin embargo, estas escenas tienen un carácter mucho más contingente, y aunque se convierten el objeto de tareas de razonamiento específico, su

retención y acceso en el largo plazo es muy costosa, y es posible que su recuerdo involucre una abstracción muy profunda y que la mayor parte de la información concreta se deseche.

Esta distinción tiene una larga historia en la representación del conocimiento, en particular en el sistema *KL-ONE* [19] que es un antecedente directo de las lógicas descriptivas [20], en el que se introdujo la distinción entre el bloque terminológico o *T-Box*, en que corresponde a la ontología en la Figura 1.4, y el bloque asertivo o *A-Box* que corresponde al conocimiento factual de la presente historia. La distinción no es trivial ya que estas dos formas de conocimiento interactúan de formas muy complejas. Por ejemplo, no es claro si los objetos individuales representados por los cuadrados se deban incluir en el bloque terminológico o el asertivo. Hay muchas conceptualizaciones posibles, cada una con sus conflictos y compromisos de expresividad particulares.

El estudio de los estados, procesos, acciones y eventos ha sido objeto de un gran esfuerzo de investigación en representación del conocimiento y se han propuesto una gran variedad de enfoques, representaciones y mecanismos de interpretación, pero hay todavía muchas preguntas abiertas y es un reto de investigación en el largo plazo.

1.7. Representación Computacional

La representación diagramática de la ontología se interpreta directamente por los seres humanos. El medio de la representación es el papel o la pantalla del monitor, y sobre éste se expresan símbolos geométricos y lingüísticos en arreglos convencionales sobre el espacio bidimensional. Sin embargo, la representación computacional del conocimiento requiere que la información se exprese en la memoria de la máquina, la cual se constituye por registros con una dirección (*Random Access Memory* o *RAM*), donde cada registro consiste a su vez de una secuencia de bytes, por ejemplo de 32 o 64, y cada byte contiene ocho bits o unidades binarias de información. Cada uno de los bytes codifica un símbolo o carácter alfabético o numérico, y el conjunto de símbolos constituye el alfabeto de la máquina. Consecuentemente la información se expresa computacionalmente como secuencias de símbolos o cadenas de

caracteres. Si la palabra, frase o oración excede el tamaño del registro la representación se expresa concatenando un número arbitrario de registros.

La traducción de la representación diagramática a su representación computacional es un proceso no trivial que involucra a la visión computacional y el procesamiento del lenguaje, y toma en cuenta las convenciones de interpretación, el contexto o dominio de conocimiento, y factores pragmáticos y de sentido común que se emplean en la comunicación. Sin embargo, por el momento asumimos que este proceso se lleva a cabo por el propio ser humano. La tarea consiste, por ejemplo, en traducir el diagrama en la Figura 1.4 a una secuencia de caracteres cuya interpretación en forma mecánica por la máquina computacional corresponda a las convenciones de interpretación expuestas a lo largo de este capítulo.

Para este efecto es necesario definir una sintaxis y una semántica. La sintaxis determina la estructura de las cadenas bien-formadas o legales que representan a las clases e individuos, y la semántica especifica las reglas o algoritmos que llevan a cabo el proceso de interpretación o evaluación.

1.7.1. Sintaxis

Una KB es una lista de forma $KB = [c_1, \dots, c_n]$ en la que cada elemento c_i es una cláusula de forma $clase(Nombre, Madre, Propiedades, Relaciones, Instancias)$ y representa a una clase. *Nombre* y *Madre* son constantes o átomos que nombran a las respectivas clases, y *Propiedades*, *Relaciones* e *Instancias* son listas que incluyen las representaciones de los objetos respectivos.

Las listas tienen forma $[e_1, \dots, e_n]$ donde los paréntesis rectangulares “[”, y “]” representan el inicio y el fin de la lista, los elementos e_i se incluyen separados por el símbolo “,” y el término “[]” representa a la lista sin elementos o vacía; asimismo, $[H|T]$ se interpreta como la lista formada por el elemento H seguido de la lista T , de acuerdo a la notación del lenguaje de programación Prolog.

El símbolo $=>$ se utiliza para asociar una propiedad con su valor (e.j., *tamaño=>grande*) o una relación con su objeto (e.j., *amigo=>pedro*). Se incluye también el símbolo $=>>$ que se

interpreta como el condicional o la implicación material así como el símbolo de igualdad = para afirmar que nombres o variables diferentes se refieren al mismo individuo.

Los términos en las listas de propiedades, relaciones e individuos se especifica con los símbolos “::=”, “;” que se interpretan como “se substituye por” y “alternativa” respectivamente, de acuerdo al uso de la notación tradicional Backus-Naur, aunque las listas se especifican siguiendo la notación de Prolog. Las reglas de formación se especifican en 1. Una *KB* es un objeto de tipo *Clases*.

```

1. átomo ::= constante ; no(constante)
2. variable ::= constante
3. peso ::= numeral
4. Predicados ::= [] ; [Predicado | Predicados]
5. Predicado ::= átomo => variable] ; variable = variable
6. Propiedades ::= [] ; [Propiedad | Propiedades]
7. Propiedad ::= [átomo, peso] ;
               [átomo => átomo, peso] ;
               [Predicados ==> Predicados, peso]
8. Relaciones ::= [] ; [Relación | Relaciones]
9. Relación ::= [átomo => átomo, peso]
10. Individuos ::= [] ; [Individuo | Individuos]
11. Individuo ::= [id => átomo ; _, Propiedades, Relaciones]
12. Clases ::= [] ; [Clase | Clases]
13. Clase ::= clase(átomo, átomo, Propiedades, Relaciones, Individuos)

```

Listado 1: Notación Base de Conocimiento No-Monotónica

Las reglas sintácticas (1-13) especifican cómo se declaran las cadenas que representan a la ontología. Las constantes son cadenas de caracteres arbitrarias que permiten nombrar una clase, una propiedad, una relación o un individuo. (1) especifica los átomos que ocurren en la KB, los cuales se pueden marcar con el prefijo o functor *no* para indicar que no se tiene la propiedad, relación o incluso la preferencia nombrada por la constante. (2) indica que se pueden expresar variables que se nombran también por constantes arbitrarias y (3) indica que

las propiedades y relaciones tienen asociado una preferencia, cuyo valor es más bajo mientras sea más prioritaria. El número 0 se utiliza para las propiedades y relaciones básicas, que tienen un carácter de necesidad en oposición a las preferencias que tienen un valor mayor que cero e indica su nivel o grado de contingencias. El peso puede variar dinámicamente ya sea por especificación directa o mediante un algoritmo de aprendizaje de máquina.

En (4-5) se especifica la forma de los predicados; éstos se utilizan para especificar los defaults condicionales en conjunto con el símbolo $\Rightarrow$ y el signo de igualdad; por ejemplo *trabaja*(*y*) y *y* = *z*. Las propiedades se especifican en (6-7) como una lista, la cual puede ser vacía. Hay tres formas: i) las propiedades booleanas se especifican por la lista de una constante asociada y su peso (e.j., [*listo*, 0]); ii) la lista de la asociación de una propiedad con su valor y su peso (e.j., [*tamaño* $\Rightarrow$ *grande*, 0]); y iii) la lista de una propiedad o default condicional y su peso de forma [*Antecedente* $\Rightarrow$ *Consecuente*, *peso*], donde *Antecedente* y *Consecuente* son listas de predicados (e.j., [[*nació*(*y*)] $\Rightarrow$ [*vive*(*y*)], 2]) y [[*vive*(*y*), *vive*(*z*)] $\Rightarrow$ [*y* = *z*], 0]). Las preferencias se incluyen en la lista de relaciones para enfatizar la intuición que las propiedades en el consecuente están condicionadas a que se tengan las propiedades en el antecedente. Esta es una convención y se podría haber incluido a las preferencias en la lista de relaciones o como un argumento adicional en las cláusulas de las clases. En (8-9) se especifica la lista de relaciones y en (10-11) la lista de individuos. Las propiedades y relaciones de las clases se leen de manera genérica y las de los individuos de manera específica. En (11) se indica que los individuos tienen un identificador pero éste se puede sub-especificar con el símbolo $_$. Se adopta asimismo la convención que las propiedades y relaciones se especifican dentro del contexto o alcance de las clases e individuos, quienes son el sujeto de la propiedad o de la relación respectivamente. Las reglas (12-13) especifican a la KB como una lista de clases. Se incluyen los átomos designados *raíz* y *ninguna* para nombrar al nodo superior de la ontología y que éste no tiene madre.

```
[
  clase(raíz, ninguna, [], [], []),
  clase(animales, raíz, [[trabaja=>y]==>[vive=>y],1],
        [[nació=>y]==>[nació=>y],2],
        [[vive=>y, vive=>z]==>[y=z],0]
        ], [], []),
  clase(aves, animales, [[vuelan,0]], [], []),
  clase(peces, animales, [], [], []),
  clase(mamíferos, animales, [[no(ovíparos),0]], [], []),
  clase(águilas, aves, [],
        [[comen=>peces,0]],
        [[id=>pedro, [[tamaño=>grande,0]],
                    [no(amigo=>arturo),0]]]]),
  clase(pingüinos, aves, [[no(vuelan),0], [color=>_,0]],
        [],
        [[id=>arturo,
          [[listo,0],
           [nació => patagonia,0],
           [trabaja => chapultepec,0]],
          [amigo=>pedro,0]]]]),
  clase(ornitorrincos, mamíferos, [[ovíparos,0]], [], [])
]
```

Listado 2: Representación de la Ontología

Por ejemplo, la ontología en la Figura 1.4 se representa en el Listado 2.¹ Se puede apreciar que la jerarquía se determina exclusivamente por los nombres de las clases y sus respectivas madres, y el orden en que se listan las cláusulas no tiene efectos en la estructura.

¹Se incluyen acentos y diéresis de acuerdo con la sintaxis española, pero éstos se deben omitir en la especificación formal en el lenguaje de programación Prolog.

1.7.2. Semántica

La semántica consiste en la interpretación de las cadenas de caracteres conforme a las reglas sintácticas. Se distinguen dos casos básicos: 1) la interpretación de los símbolos atómicos y 2) la interpretación de las expresiones compuestas. Las constantes que nombran a las clases, individuos, propiedades y relaciones se interpretan directamente. El dominio de interpretación consiste en los individuos en la ontología, las clases son las particiones respectivas, las propiedades con sus respectivos valores son los subconjuntos de individuos que tienen dicha propiedad y valor, y las relaciones los conjuntos de individuos que están en dicha relación. Las propiedades y relaciones de clases se interpretan de manera genérica; es decir, todos los individuos en la clases tienen la propiedad, o están en relación con los individuos de la clase relacionada, sin especificar quién o quienes. Por su parte, las propiedades y relaciones de individuos se interpretan de manera específica, y tienen propiedades o relaciones concretas. En esta interpretación no se permiten relaciones entre individuos particulares y clases y vice versa.

Las expresiones compuestas se interpretan de acuerdo a las convenciones ya especificadas. Estas nociones se hacen operativas determinado la *Extensión* de la KB. Ésta consiste en el dominio de interpretación incluyendo los individuos con sus respectivas clases, así como las proposiciones atómicas en la cerradura de la relación de herencia; es decir todas las proposiciones atómicas que se expresan en la base de conocimiento, ya sea explícita o implícitamente.

De forma operativa la extensión de la KB se produce mediante ocho predicados; cuatro computan la extensión o el conjunto de individuos que están en una clase, tienen una propiedad, entran a una relación como sujeto y justifican a una proposición bajo la cerradura de la relación de herencia respectivamente; y los otros cuatro se centran en un individuo y computan sus clases, propiedades, las relaciones en que entra como sujeto, y las justificaciones de sus propiedades condicionales, también bajo la cerradura de la relación de herencia. Estos predicados se especifican como sigue:

1. `extension_clase( $\alpha$ , Extension)`: provee la extensión o conjunto de individuos de la clase α bajo la relación de herencia, donde *Extension* es una variable cuyo va-

lor es la lista de individuos en la clase α . En particular `extension_clase(raíz, Extension)` provee el conjunto de todos los individuos en la KB.

2. `extension_propiedad( $\alpha$ , Extension)`: provee la extensión o conjunto de individuos que tienen la propiedad α bajo la relación de herencia, donde *Extension* es una variable cuyo valor es la lista de individuos que tienen la propiedad α –de forma *propiedad* o *propiedad* => *valor*.
3. `extension_relacion( $\alpha$ , Extension)`: provee la extensión o conjunto de individuos en la relación α bajo la relación de herencia, donde *Extension* es una variable cuyo valor es la lista de individuos que entrán como sujetos en la relacion α .
4. `extension_explicacion( $\alpha$ , Extension)`: provee la extensión o conjunto de individuos que tiene una propiedad o relación que se justifica en la propiedad o relación α bajo la relación de herencia, donde *Extension* es una variable cuyo valor es la lista de individuos
5. `clases_de( $\beta$ , Clases)`: provee el conjunto de clases del individuo α bajo la cerradura de la relación de herencias.
6. `propiedades_de( $\beta$ , Propiedades)`: provee el conjunto de propiedades del individuo o la clase α bajo la cerradura de la relación de herencias.
7. `relaciones_de( $\beta$ , Relaciones)`: provee el conjunto de relaciones en la que el individuo individuo o la clase α participa como sujeto bajo la cerradura de la relación de herencias.
8. `explicación_de( $\beta$ , Explicaciones)`: provee el conjunto de explicaciones que justifican las propiedades o relaciones que tiene el individuo α bajo la cerradura de la relación de herencias.

La extensión o el conocimiento representado el la KB resulta de evaluar estos predicados para todos las clases, propiedades, relaciones y explicaciones, así como para todos los individuos. El uso de estos servicios se hace en relación a la ontología codificada directamente en

un archivo de texto. A continuación se muestran unos ejemplos en la codificación en Prolog de estos servicios.² La evaluación se hace en relación a la ontología de la Figura 1.4, la cual se codifica textualmente en el archivo `animales.txt`. El predicado `abre_kb` lee el archivo de texto y asigna la lista de cláusulas que constituyen a la ontología a la variable *KB*. Esta variable incluye a toda la base de conocimiento y se pasa como argumento a los servicios del sistema de conocimiento. El tercer argumento de los servicios contiene el resultado de la consulta o el producto de la inferencia conceptual correspondiente.

1. Extensión de una clase:

```
% Individuos en la clase "aves"
?- abre_kb('animales.txt', KB),
   extension_clase(aves, KB, Extension_Clase).
Extension_Clase = [pedro, arturo].
```

2. Extension de una propiedad:

```
% Individuos que tienen la propiedad "nacio"
?- abre_kb('animales.txt', KB),
   extension_propiedad(nacio, KB, Extension_Propiedad).
Extension_Propiedad = [arturo:patagonia].
```

3. Extensión de una relación:

```
% Individuos que están como sujetos en la relación "no(amigo)"
?- abre_kb('animales.txt', KB),
   extension_relacion(no(amigo), KB, Extension_Relacion).
Extension_Relacion = [pedro:[arturo]].
```

4. Extension de una explicación:

```
% Justificación de porque individuos tienen la propiedad "vive=>chapultepec".
?- abre_kb('animales.txt', KB),
   extension_explicacion(vive=>chapultepec, KB, Extension_Explicacion).
Extension_Explicacion = [arturo:[trabaja=>chapultepec]].
```

5. Clases de un Individuo:

²El código de Prolog completo de todos los servicios de la base de conocimiento está disponible en <http://golem.iimas.unam.mx/lightkb>.

```
% Clases de "arturo"
?- abre_kb('animales.txt', KB),
   clases_de(arturo, KB, Clases).
Clases = [pinguinos, aves, animales, raiz].
```

6. Propiedades de un individuo:

```
% Propiedades de "arturo"
?- abre_kb('animales.txt', KB),
   propiedades_de(arturo, KB, Propiedades).
Propiedades = [listo, nacio=>patagonia, trabaja=>chapultepec, nada, no(vuela)].
```

7. Relaciones de un individuo:

```
% Relaciones de "pedro"
?- abre_kb('animales.txt', KB),
   relaciones_de(pedro, KB, Relaciones).
Relaciones = [no(amigo)=>[arturo], comen=>[peces]].
```

8. Explicaciones relativas a un individuo

```
% Explicaciones para "arturo"
?- abre_kb('animales.txt', KB),
   explicaciones_de(arturo, KB, Explicaciones).
Explicaciones = [[vive=>chapultepec]:[trabaja=>chapultepec]].
```

Los predicados (1-8) se especifican con algoritmos que recorren la estructura de la taxonomía. Estas búsquedas admiten diversas implementaciones, por ejemplo desde la raíz hasta las hojas (*top-down*) o de las hojas hacia la raíz (*bottom-up*) así como diversos órdenes de recorrido, por ejemplo, por búsqueda de profundidad por la izquierda o por la derecha, pero la estrategia es independiente del resultado de la consulta que depende de factores estructurales exclusivamente. Se debe considerar que algunas búsquedas se centran en objetos que sólo aparecen en una región del dominio, como los individuos y las clases, en cuyo caso sólo se requiere visitar los segmentos correspondientes de la jerarquía, o en objetos que pueden aparecer en cualquier región, como las propiedades y las relaciones, ya sean absolutas o condicionadas, pero en todo caso los algoritmos deben especificarse de forma declarativa para verificar que produzcan resultados válidos, así como su complejidad, para verificar que se

pueden computar con recursos finitos de memoria y tiempo de procesamiento. El desarrollo y análisis de los algoritmos es un tema abierto de investigación.

La implementación de la negación, el razonamiento con conocimiento incompleto y el principio de especificidad requiere que la información se colecte de acuerdo a una convención estipulada explícitamente. Por ejemplo, que las listas que representan a la extensión de la base de conocimiento retribuida por cada uno de los ocho servicios incluyan la información más específica o más prioritaria al frente. Los algoritmos recopilan todos los átomos, positivos y negativos, que se encuentran a lo largo del recorrido de la región de referencias, pero la respuesta final de la consulta incluye sólo la opción preferida. Hay tres casos: i) si el objeto de referencia es un átomo positivo y aparece al frente de la lista, la respuesta es que el objeto tiene la propiedad o relación objeto de la pregunta; ii) si el objeto de referencia es un átomo negativo y así aparece al frente de la lista no la tiene y iii) si en la lista no se encuentra la forma positiva ni la negativa, la respuesta del sistema de conocimiento es *no sé*. Si *Termino* es una consulta al servicio de conocimiento, es decir un átomo que nombra a una propiedad, a una propiedad con su valor o a una relación, y el tercer argumento de los servicios (1-8) se designa de manera genérica por la variable de Prolog *Extension* la respuesta final producida por el servicio correspondiente se obtiene explorando el valor de dicha variable, es decir la extensión del tipo de objeto requerido, de izquierda a derecha, por el procedimiento recursivo en el Listado 3. Una implementación particular de los ocho servicios de la base de conocimiento, incluyendo los procedimientos para resolver los conflictos de conocimiento mediante el principio de especificidad y los pesos de las preferencias, está disponible en <http://golem.iimas.unam.mx/lightkb>.

Por ejemplo, la respuesta a la pregunta *tiene un Individuo la Propiedad?* en relación a una ontología KB se responde mediante el procedimiento en el Listado 4:

Los resultados de estos algoritmos se pueden conceptualizar como el producto de las inferencias conceptuales que realizan la recuperación de la información almacenada en la memoria de largo plazo, tanto de la información registrada de forma explícita como implícita. La extensión total o parcial de la KB que se produce mediante los servicios del sistema de conocimiento puede ser objeto de preguntas específicas formuladas por medio de programas de

```

% El átomo no aparece en forma positiva ni negativa
conocimiento_incompleto(_, [], 'no se').

% El átomo es negativo
conocimiento_incompleto(not (Exp), [not (Exp) | _], not (Exp)).

% El átomo es positivo
conocimiento_incompleto(Exp, [Exp | _], Exp).

% El átomo no aparece al frente de la lista
conocimiento_incompleto(Exp, [_ | Resto], Respuesta) :-
    conocimiento_incompleto(Exp, Resto, Respuesta).

```

Listado 3: Interpretation of the KB with Incomplete Information

```

propiedad_individuo(Individuo, Propiedad, KB, Resp) :-
    propiedades_de(Individuo, KB, Props),
    conocimiento_incompleto(Propiedad, Props, Resp).

```

Listado 4: Consulta con conocimiento incompleto

Prolog estándar, que corresponden al proceso de la memoria de trabajo. En general, aplicaciones específicas requieren la definición de un conjunto de preguntas genéricas que se definen de antemano, y se utilizan de manera cotidiana en los sistemas de conocimiento. El uso de un ambiente de programación declarativo y flexible como Prolog permite adicionalmente la definición de preguntas particulares que se puedan anticipar.

El sistema de representación incluye adicionalmente servicios para agregar, modificar y eliminar objetos de conocimiento de todos los tipos de forma dinámica, incluyendo clases e individuos, con sus propiedades y valores, relaciones y preferencias. La actualización puede incluir proposiciones incoherentes e incluso contradictorias en relación al contenido de

la memoria y, sin embargo, los servicios de inferencia deben proveer de la respuesta correcta relativa a las convenciones de interpretación en el nuevo estado. En particular, aumentar o remover una clase consiste en crear o remover una partición del dominio sin afectar a los individuos en la región modificada, aunque éstos cambian de clase y se incluyen en la que corresponda a la nueva partición, la cual puede ser más específica o más genérica según se agreguen o se remuevan particiones respectivamente. Estas operaciones permiten cambiar el modelo conceptual de la base de conocimiento de manera incremental. Los individuos con sus propiedades y relaciones se pueden actualizar directamente, y aumentar, eliminar o cambiar el valor de propiedades y relaciones, ya sea de clases o individuos, se lleva a cabo mediante operaciones sobre las listas en que se representan dichos objetos. La implementación en Prolog del conjunto de servicios del presente sistema de conocimiento está también disponible en <http://golem.iimas.unam.mx/lightkb>.

1.8. Memoria e Inferencia

La representación de los conceptos que constituyen al conocimiento simbólico se almacena en la memoria de largo plazo. Ésta incluye dos categorías principales: la semántica y la episódica [21]. La primera incluye el conocimiento lingüístico y terminológico, el enciclopédico, el que se tiene por pertenecer a una comunidad lingüística y cultural, entre muchos otros tipos posibles, y el segundo la historia de vida o autobiografía del propio agente. Los seres humanos somos capaces de almacenar todo este conocimiento y de retribuirlo de manera muy efectiva cuando se requiere en la experiencia cotidiana y, en particular, en el uso del lenguaje y el ejercicio del pensamiento.

La taxonomía de los animales ilustra que el conocimiento es un objeto muy complejo. La extensión de la base de conocimiento en la Figura 1.4 es limitada pero si se incluyera todo el conocimiento que se tiene acerca de los animales, y más aún, el conocimiento de la vida cotidiana de un ser humano, sería de dimensiones extraordinarias. El problema es mucho mayor si se considera que el conocimiento humano es no monotónico y que el incremento de la capacidad expresiva conlleva la introducción de diversos tipos de coherencias y contra-

dicciones; sin embargo, los seres humanos tenemos la capacidad de razonar productivamente de manera muy efectiva.

Las contradicciones surgen del conflicto entre la estructura de la representación y los medios expresivos. La estructura jerárquica interactúa con la negación y la implicación lógica, y la interpretación se requiere complementar con el principio de especificidad y el uso de preferencias. Estos conflictos son contingentes a la estructura jerárquica –en particular, la taxonomía no monotónica y con preferencias presentada aquí fue desarrollada en el contexto del proyecto Golem, y es la que se usa en el robot Golem-III [22, 23, 24], y se ha aplicado también al desarrollo de sistemas de información no regimentados [25]– pero toda estructura representacional tiene que enfrentar sus conflictos particulares. Otro formalismo no monotónico con una orientación lógica y semántica es la llamada *Answer Set Programming* [26].

Otros esquemas, aunque no necesariamente no monotónicos, son los Marcos de Minsky, las redes semánticas, en sus diversas variedades y estilos; las primitivas conceptuales, o las representaciones en lenguaje formales, lógicos y funcionales. Otro enfoque es el de las llamadas lógicas descriptivas, que distinguen la información terminológica de la memoria semántica, en la llamada T-Box, de la información factual, que corresponde a la memoria episódica o asertiva, en la llamada A-Box. Otro más es el de los llamados sistemas expertos basado en reglas de producción, muy cercano al programa inicial de la racionalidad limitada. Esta variedad es muy amplia y el desarrollo y aplicación de esquemas representacionales es el pan de cada día de la Inteligencia Artificial.

A pesar de el esfuerzo de investigación muy intenso que se ha realizado desde el inicio de la Inteligencia Artificial con el artículo de Turing de 1950, todavía no se sabe cuál es la estructura de la memoria humana ni cuáles son sus recursos expresivos, ni qué conflictos de conocimiento surgen y cómo se resuelven. Es muy intuitivo que tanto en la memoria artificial como en las memorias naturales se tiene que abordar el registro, el reconocimiento y la retribución de la información, y cualquier modelo se tiene que referir a un formato o estructura particular, pero la creación de un modelo capaz de representar al conocimiento conceptual humano es un problema de investigación abierto.

En particular todo sistema de representación enfrenta el llamado compromiso de la representación del conocimiento (*the knowledge representation trade-off*) [27, 28] que establece que los recursos computacionales que se requieren para interpretar una representación varían de forma inversa a su expresividad. Interpretar representaciones con información concreta es fácil pero interpretar abstracciones es costoso, y el esquema de representación debe ser suficientemente expresivo para capturar al dominio, pero al mismo tiempo, lo menos expresivo posible para que se pueda interpretar con recursos computacionales limitados.

El compromiso de la representación del conocimiento se aplica a todos los sistemas simbólicos artificiales, como los que se pueden representar en computadoras digitales del tipo ordinario; sin embargo, no es claro si se aplica a memorias naturales ya que es muy intuitivo que razonar con información concreta es fácil si la información es limitada, pero si ésta aumenta, o el dominio de conocimiento es infinito, como en el caso de los objetos matemáticos, se tienen que introducir abstracciones para razonar de manera efectiva, y se sigue que las abstracciones se forman precisamente para razonar fácilmente con estructuras complejas o con grandes cantidades de información. Esta oposición entre la computación simbólica y el razonamiento natural es una paradoja de la representación del conocimiento [11].

En la inferencia conceptual “se piensa en la memoria”. Si se pregunta si las águilas son animales carnívoros o si los ornitorrincos ponen huevos, las respuestas se presentan inmediatamente a la consciencia. Las operaciones de la memoria no son accesibles a la introspección, de manera muy similar a las informaciones que se presentan a la consciencia a partir de la percepción.