Universidad de Buenos Aires
Facultad de Ciencias Exactas y Naturales
Departamento de Computación

Bisimulaciones en Neighbourhood Semantics

Tesis presentada para optar al título de
Licenciado en Ciencias de la Computación
de la Universidad de Buenos Aires

Diego Federico Figueira

Director de tesis: Dr. Carlos Areces

Lugar de trabajo: Loria, Vandœuvre-lès-Nancy, Francia

Buenos Aires, 2006

A Nadis

ÍNDICE

BISIMULACIONES EN NEIGHBOURHOOD SEMANTICS

Resumen. En esta tesis investigamos la semántica de lógica modal conocida como Neighbourhood Semantics. Existen dos definiciones alternativas que resultan en lógicas distintas, la primera de las cuales fue introducida por Montague y Scott en los años '70 ($\mathcal{N}_=$), y la segunda por van Benthem años más tarde ($\mathcal{N}_\subseteq$).

Estudiamos y definimos la noción de bisimulación en ambas lógicas, siendo que $\mathcal{N}_=$ resulta tener una bisimulación compleja. Para esta lógica analizamos dos extensiones buscando en cada caso la bisimulación correspondiente. Llegaremos a una noción de bisimulación intuitiva para una extensión clásica con la modalidad universal: $\mathcal{N}_=(\mathsf{E})$.

En la parte final del trabajo, estudiamos la complejidad computacional del problema de satisfacibilidad (SAT) en cada lógica. Mostramos que SAT-$\mathcal{N}_\subseteq$ está en *NP* (y es por lo tanto *NP-Complete*), y que SAT-$\mathcal{N}_=(\mathsf{E})$ está en *EXP-Time*.

Palabras clave. Neighbourhood Semantics, lógica epistémica, estructuras Montague-Scott, modelo minimal, bisimulación, esquemas M, C y N, complejidad computacional.

BISIMULATIONS ON NEIGHBOURHOOD SEMANTICS

Summary. In this thesis we investigate the modal logic semantics known as Neighbourhood Semantics. There are two alternative definitions that correspond to different logics, the first of which was introduced by Montague and Scott in 1970 ($\mathcal{N}_=$), and the second by van Benthem some years later ($\mathcal{N}_\subseteq$).

We study and define the notion of bisimulation in both logics, and observe that $\mathcal{N}_=$ has a complex bisimulation. For this logic we analyse two extensions defining in each case the corresponding bisimulation. We will meet a bisimulation which is intuitive for a classical extension of the basic language with the universal modality: $\mathcal{N}_=(\mathsf{E})$.

In the final part of this work, we study the computational complexity of the satisfiability problem (SAT) of each logic. We show that SAT-$\mathcal{N}_\subseteq$ is in *NP* (and therefore *NP-Complete*), and that SAT-$\mathcal{N}_=(\mathsf{E})$ has an *EXP-Time* upper-bound.

Key words. Neighbourhood Semantics, epistemic logic, Montague-Scott structures, minimal model, bisimulation, schemas M, C and N, computational complexity.

AGRADECIMIENTOS

En primer lugar, quisiera expresar mi profundo agradecimiento a mi director de tesis Carlos Areces por su paciencia, disponibilidad y hospitalidad en todo momento. Agradezco especialmente a Patrick Blackburn, director del grupo de investigación *Langue et Dialogue* (LeD) dentro de Loria, por apoyar este trabajo, y por sus comentarios y críticas tan enriquecedoras.

Todos los resultados de esta tesis fueron relizados durante mi estadía en Francia dentro del laboratorio Loria (*Laboratoire Lorrain de Recherche en Informatique et ses Applications*) gracias a una beca concedida por Inria (*Institut National de Recherche en Informatique et en Automatique*). Por tener esa oportunidad, estoy agradecido a Inria International.

A mis compañeros de estadía en Francia, Daniel Gorín y Sergio Mera, por crear una atmósfera agradable y cómoda.

Al Grupo de Lógica y Computabilidad (GLyC) por interesarse en mi investigación y brindarme todo el apoyo técnico que necesité, en particular a mi hermano Santiago Figueira.

A mi familia y amigos.

ACERCA DE ESTA TESIS

Enfoque

Las lógicas modales son lógicas usualmente decidibles y más expresivas que la lógica proposicional, que fueron usadas originalmente para capturar nociones como *necesidad* u *obligación* y razonar sobre ellas. La semántica más usual para estas lógicas es la *relacional*, definida en términos de estructuras relacionales como las usadas para interpretar un lenguaje de primer orden con una signatura sin símbolos de función. Un modelo de esta semántica puede ser visto como un grafo dirigido en donde dos elementos se representan conectados mediante un arco cada vez que están relacionados mediante alguna relación del modelo (llamadas *relaciones de accesibilidad*). Bajo esta semántica, las lógicas modales pueden verse como lenguajes de primer orden con cuantificación acotada a elementos *accesibles*. La historia de la lógica modal ha probado lo útil y versátil que es esta semántica para capturar muchos tipos de operadores lógicos, incluyendo nociones temporales (futuro, pasado, ahora), predicamentos universales (en todo posible estado), etc.

Sin embargo, esta semántica es inadecuada para razonar sobre algunas nociones clásicas de las lógicas epistémicas como ser la noción de *creencia* o *conocimiento*. La semántica relacional trae aparejada algunas inferencias que resultan incorrectas con estas nociones epistémicas. Esto motivó a Montague y Scott a proponer una semántica alternativa. Dicha semántica se conoce como *Neighbourhood Semantics*, y en esta tesis la llamaremos *Neighbourhood Semantics Strict*, ya que como veremos, más tarde aparecieron propuestas alternativas.

Una de las herramientas más conocidas en lógica modal que revela el poder expresivo y condiciones para la invarianza semántica de una lógica modal, es la de *bisimulación*. Esta noción fue introducida por Johan van Benthem (inicialmente con el nombre de *p-relación*) en su tesis doctoral de 1976 [12]. Bajo esta lupa de bisimulación, la semántica de *Neighbourhood Semantics* introducida por Montague y Scott no resulta tan atractiva: su noción de bisimulación es inherentemente compleja.

Quizás debido a esto, años más tarde Johan van Benthem introdujo, bajo el mismo nombre de *Neighbourhood Semantics* (en el presente trabajo la llamaremos *Neighbourhood Semantics Loose*), otra lógica modal que también resultaba útil para manejar algunas nociones epistémicas pero cuya definición en términos de bisimulación resultaba más directa y natural.

El presente trabajo se centrará en discutir el poder expresivo de estas dos definiciones de *Neighbourhood Semantics*. Nos enfocaremos en dos formas de explorar el poder expresivo. Por un lado estudiaremos la noción de bisimulación que identifica a cada lógica. Para esto definiremos las nociones de bisimulación en ambas lógicas y, en el caso de *Neighbourhood Semantics Strict*, buscaremos una extensión de la lógica que resulte en una noción simplificada de bisimulación. Por otro lado, investigaremos la complejidad computacional del problema de satisfacibilidad en cada lógica, para comprender la dificultad de cálculo a la que nos enfrentamos.

Lo que vendrá

En el capítulo 1 de esta tesis se introducirán las lógicas modales; describiremos las nociones básicas para su comprensión y daremos ejemplos de algunas de sus variedades. Esto lo haremos utilizando la noción clásica de semántica relacional. Por último se presentará la noción de bisimulación y se discutirá su importancia en la lógica modal.

Dedicaremos el capítulo 2 a exponer la necesidad de una semántica más general que la relacional. Mostraremos que ambas lógicas definidas en términos de *Neighbourhood Semantics* suplen esta necesidad, y estableceremos las diferencias entre éstas y la lógica relacional.

Para estudiar el poder expresivo de *Neighbourhood Semantics*, destinaremos el capítulo 3 a discutir las nociones de bisimulación para ambas lógicas. Mostraremos las dificultades que aparecen al intentar definir la noción apropiada de bisimulación para la lógica *Neighbourhood Semantics Strict*. A continuación mostraremos dos extensiones de esta lógica para llegar a una bisimulación más intuitiva.

Continuando con el estudio del poder expresivo de ambas lógicas, pero esta vez desde una perspectiva computacional, en el capítulo 4 haremos una comparación entre estas dos lógicas, mostrando sus similitudes y buscando cotas para la complejidad del problema de satisfacibilidad.

1. INTRODUCCIÓN

1.1 Primeros pasos de la lógica modal

Podemos ubicar el origen de las lógicas modales como disciplina matemática en 1918, cuando C. I. Lewis publica su *Survey of Symbolic Logic* [7]. Si bien existen trabajos relacionados con lógicas que podríamos considerar *modales* desde la época de Aristóteles hasta bien entrado el siglo XIX, no es fácil relacionarlos con la tradición de la lógica moderna.

En [7], Lewis analiza la diferencia que hay entre implicaciones de la forma $A \to B$ donde el hecho de que B no es falso cuando A es verdadero es *contingente* y aquellas donde este hecho es *necesario.* Alguien puede decir "si llueve, entonces Juan no va a trabajar" y esta afirmación puede ser *verdadera* en el contexto de enunciación (Juan puede estar engripado, o vivir en una zona inundable o vender bronceador en la playa), pero claramente no se trata de una *verdad universal* (dejará de ser verdadero si Juan se cura, se muda o cambia de trabajo). La relación que hay entre "que llueva" y "que Juan no vaya a trabajar" es contingente. Por otro lado, entre las sentencias "si llueve, Juan no va a trabajar" y "como Juan va a trabajar, entonces no llueve" hay una relación lógica más fuerte (*logical entailment*) de forma que en cualquier contexto en que la primera se cumpla, la segunda necesariamente también debe valer.

En sus trabajos, Lewis extiende la lógica proposicional agregando el *operador modal* $\Diamond$, que se interpreta como "es posible que" y con él define el operador de *implicación estricta* $\mapsto$ de la siguiente forma: $\varphi \mapsto \psi \equiv \neg\Diamond(\varphi \wedge \neg\psi)$; esta última fórmula se debe leer como "no es posible (en ningún contexto imaginable) que simultáneamente φ sea verdadero y ψ no lo sea". Usando este lenguaje podemos decir "si llueve, Juan no va a trabajar" y "Juan va a trabajar, entonces no llueve" utilizando la implicación estricta: $(\mathsf{llueve} \to \neg\mathsf{trabaja}_{\mathsf{juan}}) \mapsto (\mathsf{trabaja}_{\mathsf{juan}} \to \neg\mathsf{llueve})$ y estaremos afirmando que éste es un hecho que no depende de circunstancias particulares.

$\Diamond$ es un operador modal ya que asigna un *modo* a un valor de verdad. Si $\Diamond$ toma el modo de *posibilidad*, lo usamos para diferenciar lo que *es verdadero* (o falso) de lo que *podría ser verdadero* (o falso). Tomemos dos proposiciones: $\mathsf{nublado}$ y llueve, y veamos ejemplos de lo que podemos decir al asociar este modo con dicho operador:

Fórmula	Significado
$\mathsf{nublado}$	Está nublado en este momento
llueve	Está lloviendo en este momento
$\Diamond\mathsf{nublado}$	Podría estar nublado en este momento
$\neg\Diamond\mathsf{llueve}$	No podría llover en este momento
$\neg\mathsf{nublado} \wedge \mathsf{llueve}$	En este momento llueve aunque no está nublado
$\neg\Diamond(\neg\mathsf{nublado} \wedge \mathsf{llueve})$	No podría suceder que ahora esté lloviendo aunque no esté nublado

En este ejemplo vemos dos casos de la forma "no es posible que suceda X". Intuitivamente, esto es lo mismo que decir "es necesario que no suceda X". La *necesidad* es una modalidad que se suele notar con el operador $\Box$. Como vimos, existe una relación muy fuerte entre *posibilidad* y *necesidad*: $\neg\Diamond\neg\varphi \equiv \Box\varphi$. Cuando dos operadores modales cumplen con esta propiedad, se dice que uno es el *dual* del otro. Tradicionalmente, se asume que si $\Diamond$ es una modalidad cualquiera, $\Box$ es su dual (y viceversa).

Interpretar informalmente una fórmula modal como hicimos en el ejemplo anterior no es complicado. De hecho, en ese ejemplo podríamos haberle asignado a $\Diamond$ un modo cualquiera, por ejemplo *deseo*, e interpretado $\Diamond\mathsf{nublado}$ como "desearía que estuviera nublado". Hasta aquí, lo que tenemos no es más que una forma de escribir simbólicamente algunos enunciados. Para *capturar* efectivamente una noción imprecisa como "la posibilidad", Lewis analizó qué inferencias son válidas al utilizar este modo. Por ejemplo, inferir que $\Diamond\mathsf{llueve}$ dado que llueve es correcto si $\Diamond$ representa *posibilidad* (no es admisible aceptar como válido que "llueve pero no es posible que llueva"), pero claramente no lo es si $\Diamond$ representa *deseo*. En sus trabajos, Lewis presenta diversos sistemas axiomáticos que intentan generar como teoremas todas las fórmulas que son *verdaderas* en una lógica dada y sólo dichas fórmulas (la primera característica hace a un sistema axiomático *completo*, la segunda lo hace *correcto*).

A partir de los trabajos de Lewis, surgió el interés de buscar nuevas extensiones *modales* de la lógica proposicional. Se investigaron, de esta forma, las axiomatizaciones de conceptos tales como *obligación*, *creencia*, *conocimiento*, etc. Lo que, en general, estos primeros lógicos modales hicieron fue capturar de una manera puramente sintáctica conceptos que hasta ese momento pertenecían al dominio de la intuición.

Definimos el lenguaje que llamaremos *lógica modal básica*, como una extensión del lenguaje de la lógica proposicional con la inclusión del operador modal "$\Box$".

Definición 1.1. Dado PROP un conjunto numerable de símbolos de proposición, el conjunto $\mathcal{M}$ de fórmulas bien formadas de la lógica modal básica definidas sobre PROP se define inductivamente como:

$$\mathcal{M} ::= p \mid \neg\varphi \mid \varphi \wedge \varphi' \mid \Box\varphi$$

donde $p \in \mathsf{PROP}$ y $\varphi, \varphi' \in \mathcal{M}$. El operador $\Diamond$ se define como $\Diamond\varphi \equiv \neg\Box\neg\varphi$. El resto de los operadores booleanos se definen de la forma usual.

Un trabajo que merece un comentario aparte es el que realizó Arthur Prior a principios de la década de 1950 con una familia especial de lógicas modales: las llamadas *lógicas temporales* (*temporal logics* o *tense logics*). Prior intentaba representar en una lógica la forma en la cual tratamos las relaciones temporales en los lenguajes naturales (de ahí su nombre original de *tense logics*); para ello introduce la ubicación temporal como modo. En la lógica temporal básica, Prior utiliza las modalidades $\langle F\rangle$ y $\langle P\rangle$. La interpretación que se espera de una fórmula $\langle F\rangle\varphi$ es "φ será verdadera en algún momento en el ***F***uturo.", y la interpretación de $\langle P\rangle\varphi$ es "φ fue verdadero en algún momento en el ***P***asado." Es usual escribir $\langle F\rangle$ como F y $\langle P\rangle$ como P, y sus duales se escriben como G y H respectivamente[1], y permiten predicar sobre todos los instantes del futuro o del pasado.

En este lenguaje podemos, por ejemplo, escribir la frase "siempre que llovió, paró" como $H(\mathsf{llueve} \to F(\neg\mathsf{llueve}))$ y muchas afirmaciones generales sobre el tiempo. Por ejemplo, $\varphi \to GP\varphi$, dice "cualquier cosa que suceda, siempre habrá de haber sucedido", y esto parece un candidato para una verdad universal acerca del tiempo. Por otro lado, si decimos que

[1] Estos operadores son mnemotécnicos sólo en inglés: "it is always ***G***oing to be the case" y "it always ***H***as been the case".

$F\varphi \to FF\varphi$ debe ser siempre cierto, estamos pensando en la noción del tiempo como *densa*: entre dos instantes de tiempo, siempre hay un tercero.

Notemos que, si bien F y P son operadores modales que están íntimamente relacionados (por ejemplo, $\varphi \to \neg F \neg P\varphi$ es verdadero, para todo φ), no es posible escribir uno en función del otro. Ambos deben incluirse como operadores primitivos de la lógica temporal; a partir de ellos es posible derivar G y H.

En esta lógica, si quisiéramos predicar sobre una verdad atemporal φ, que es cierta en todo instante, podríamos expresarlo como $H\varphi \land \varphi \land G\varphi$. Pero esta noción de cuantificación universal puede tener sentido en otras lógicas y trasciende la temporal.

Abstrayendo este concepto, podemos modelar una perspectiva *global* que predica sobre *todos* los contextos o mundos posibles de una lógica (en el caso de la lógica temporal, todos los *instantes*) introduciendo la *modalidad universal* (que escribiremos E para su forma de diamante y A para su forma dual) con una interpretación fija. Así, A y E son operadores que cuantifican sobre *todo* posible mundo de forma tal de interpretar la fórmula $\mathsf{E}\varphi$ como "En *algún* mundo φ es verdadera" y $\mathsf{A}\varphi$ como "En *todo* mundo φ es verdadera".

1.2 La revolución semántica

Treinta años después de la publicación del trabajo original de Lewis, el interés por las lógicas modales se estaba extinguiendo. Hasta ese momento, cada una de las lógicas modales se caracterizaba únicamente por el conjunto de sus teoremas, los cuáles a su vez se expresaban utilizando sistemas axiomáticos. Pero era, justamente, este enfoque esencialmente sintáctico que tenían las investigaciones en el área lo que frenaba su desarrollo. A modo de ejemplo, cada vez que se proponía una nueva lógica, se daba su sistema axiomático característico, y se la clasificaba buscando su relación de inclusión respecto a otras lógicas conocidas. Sin embargo, mostrar que hay una relación de inclusión o no-inclusión entre los lenguajes que inducen dos sistemas axiomáticos arbitrarios es una tarea inherentemente compleja.

Se puede decir, sin exagerar, que se vivió una verdadera *revolución* en el área cuando, a principios de la década de 1960, diversos trabajos, entre los que sobresalen los de Samuel Kripke, mostraron que era posible encarar el estudio de las lógicas modales desde un punto de vista completamente novedoso. Lo que estos investigadores descubrieron es que las lógicas modales pueden ser *interpretadas* en términos de estructuras matemáticas precisas. A partir de ese momento, una lógica modal dada pasó a caracterizarse no sólo por el conjunto de sus tautologías, sino también por el *tipo de modelos* que admitía. Tareas como la clasificación de una nueva lógica que mencionamos más arriba podían hacerse ahora *semánticamente*, comparando modelos, de una forma simple y elegante.

Estas estructuras tomadas como modelos de lógica modal se conocen hoy en día como *modelos de Kripke* (también llamadas *modelos relacionales* o *modelos de mundos posibles*). Desde un punto de vista computacional, un modelo de Kripke es un grafo dirigido donde cada nodo está etiquetado por un conjunto de símbolos proposicionales. Intuitivamente, cada nodo del grafo representa un *contexto* donde es posible evaluar una fórmula modal, la relación de accesibilidad entre nodos está asociada al operador modal y la etiqueta de cada punto es una *valuación* que permite determinar el valor de verdad de las proposiciones en dicho contexto.

Definición 1.2 (Modelo de Kripke)**.** Dado PROP un conjunto numerable de símbolos de proposición, un modelo de Kripke es una estructura $M = \langle W, R, V \rangle$ donde:

$$\begin{array}{rl} W & \text{es un conjunto no vacío,} \\ R \subseteq W \times W & \text{es una relación binaria, y} \\ V(p_i) \subseteq W & \text{para cada } p_i \in \mathsf{PROP}. \end{array}$$

La relación R se denomina la *relación de accesibilidad* entre elementos del modelo. Los elementos del modelo son llamados estados, nodos o mundos indistintamente.

Definición 1.3 (Interpretación de una fórmula de $\mathcal{M}$)**.** Dado un lenguaje modal $\mathcal{M}$ definido sobre PROP, y una estructura $M = \langle W, R, V \rangle$, la relación de satisfacibilidad $\models$ se define recursivamente como:

$$\begin{array}{lll} \mathcal{M}, w \models p_i & \text{sii} & w \in V(p_i),\ p_i \in \mathsf{PROP} \\ \mathcal{M}, w \models \neg\varphi & \text{sii} & \mathcal{M}, w \not\models \varphi \\ \mathcal{M}, w \models \varphi_1 \wedge \varphi_2 & \text{sii} & \mathcal{M}, w \models \varphi_1 \text{ y } \mathcal{M}, w \models \varphi_2 \\ \mathcal{M}, w \models \Box\varphi & \text{sii} & \text{para todo } w' \in W \text{ tal que } wRw' \text{ se verifica } \mathcal{M}, w' \models \varphi. \end{array}$$

Cuando para todo $w \in W$ se cumple $\mathcal{M}, w \models \varphi$ decimos que φ es *válida* en $\mathcal{M}$ y lo notamos $\mathcal{M} \models \varphi$. Si incluso para *cualquier* modelo $\mathcal{M}$ se cumple $\mathcal{M} \models \varphi$, decimos que φ es *universalmente válida* y lo notamos $\models \varphi$.

Por otro lado, también introduciremos la noción de *extensión* de una fórmula, que no es más que los nodos en donde dicha fórmula se verifica.

Definición 1.4 (Extensión de una fórmula)**.** Definimos la *extensión* o *significado* de una fórmula φ sobre un modelo $\mathcal{M}$ como el conjunto de mundos que la satisfacen:

$$\|\varphi\|^{\mathcal{M}} := \{w \mid \mathcal{M}, w \models \varphi\}.$$

Los modelos de Kripke resultaron útiles para dar semántica no sólo a la lógica recién presentada. De hecho, Kripke y otros descubrieron que los sistemas axiomáticos que se habían usado durante treinta años para capturar distintas nociones intuitivas como posibilidad u obligación, en realidad estaban caracterizando sintácticamente clases de modelos de interés más general. Por ejemplo, el sistema $S4$, uno de los investigados por Lewis en sus trabajos sobre *posibilidad*, genera todas las fórmulas –y sólo las fórmulas– que son *válidas* en los modelos cuya relación de accesibilidad es transitiva y reflexiva.

Sobre esta semántica relacional se puede definir la lógica temporal de forma modular sobre el mismo concepto de modelo de Kripke. Tomemos la relación R del modelo como aquella que relaciona dos estados cada vez que el primero es temporalmente anterior al segundo. De esta forma podemos definir F de la misma forma que definiríamos $\Diamond$ para la lógica modal básica:

$$\mathcal{M}, w \models F\varphi \quad \text{sii} \quad \text{existe } w' \in W \text{ tal que } wRw' \text{ y se verifica } \mathcal{M}, w' \models \varphi$$

Por otro lado, para definir P debemos navegar la relación R en dirección opuesta:

$$\mathcal{M}, w \models P\varphi \quad \text{sii} \quad \text{existe } w' \in W \text{ tal que } w'Rw \text{ y se verifica } \mathcal{M}, w' \models \varphi$$

Los operadores F y P se definen en realidad de la misma forma que $\Diamond$, y es por eso que la lógica podría definirse también en términos de dos modalidades estándar (llamémoslas $\langle F \rangle$ y $\langle P \rangle$) dentro de una lógica modal básica, cada una asociada a una relación R_1 y R_2 en donde nos debemos restringir a aquellos modelos en donde $R_1 = R_2^{-1}$. Estas dos aproximaciones resultan equivalentes y se pueden usar indistintamente.

La noción de modalidad universal también puede ser definida sobre los modelos de Kripke de la siguiente forma[2]:

$$\begin{array}{lll} \mathcal{M}, w \models \mathsf{E}\varphi & \text{sii} & \text{hay un } u \in W \text{ tal que } \mathcal{M}, u \models \varphi \\ \mathcal{M}, w \models \mathsf{A}\varphi & \text{sii} & \text{para todo } u \in W \text{ tenemos que } \mathcal{M}, u \models \varphi. \end{array}$$

Todos estos resultados atrajeron a investigadores de nuevas ramas de la matemática hacia las lógicas modales, lo cual revitalizó el trabajo del área. Sin embargo, muchos investigadores tradicionales, que se habían interesado por las connotaciones filosóficas de las distintas modalidades, lentamente fueron perdiendo interés en el tema.

Hoy en día ya no se habla de *la* lógica modal. Existe una gran de variedad de lógicas con modalidades, desde las clásicas unimodales hasta lógicas con infinitas modalidades (e.g., PDL [6]). En las últimas dos décadas las lógicas modales se han convertido en herramientas de utilidad práctica, especialmente en diversas áreas de la lingüística y la computación. Dentro de esta última, se usan lógicas modales en ámbitos tan variados como la especificación y verificación de sistemas, la representación del conocimiento, y la realización de *planning* en inteligencia artificial.

¿Por qué resultan las lógicas modales útiles en disciplinas tan diversas? Tal vez se deba a una combinación de razones. Para empezar, en muchas disciplinas se trabaja continuamente con estructuras que se pueden interpretar como grafos dirigidos y estas lógicas son ideales para expresar propiedades de este tipo de estructuras (ya vimos en la Definición 1.2 que un modelo de Kripke no es más que un grafo). Además, las modalidades son, de alguna manera, modulares; esto es, uno puede agregar exactamente aquellas modalidades que necesite y obtener así una lógica *a medida*. Por último, estas lógicas tienen en general buenas propiedades metalógicas y computacionales; en particular, la validez de una fórmula constituye un problema usualmente decidible.

1.3 Caracterización axiomática de la lógica modal básica

Si bien en esta tesis nuestro principal interés es semántico, necesitaremos trabajar también con algunas nociones sintácticas en el capítulo 4. Los interrogantes que intentamos responder ahora son, si estamos en la clase de modelos M, ¿hay mecanismos sintácticos capaces de generar Λ_{M}, las fórmulas válidas en M? Y estos mecanismos ¿son capaces de capturar la relación semántica de consecuencia asociada?

Una lógica normal es simplemente un conjunto de fórmulas que satisfacen ciertas condiciones de clausura sintáctica. Definiremos ahora un sistema axiomático à-la-Hilbert llamado **K**. **K** es el sistema "mínimo" (o más débil) para razonar acerca de modelos relacionales, y agregando axiomas adicionales se obtienen sistemas más fuertes.

Definición 1.5. Una **K**-*prueba* es una secuencia finita de fórmulas, cada una de las cuales es un *axioma*, o se sigue de elementos anteriores de la secuencia aplicando una *regla de inferencia*. Los axiomas de **K** son todas las instancias de tautologías proposicionales, y además:

[2] Notar que, dado que A y E son operadores *duales*, podríamos dar la semántica de sólo uno de ellos, y definir el restante con la equivalencia $\mathsf{A}\varphi \equiv \neg\mathsf{E}\neg\varphi$.

(K) $\Box(p \to q) \to (\Box p \to \Box q)$

La noción de *teorema* en **K** puede definirse inductivamente a partir de los axiomas y sus *reglas de inferencia*:

- Todos los axiomas son teoremas.
- *Modus ponens*: si φ y $\varphi \to \psi$ son teoremas, ψ es teorema.
- *Substitución uniforme*: si φ es teorema, θ es teorema, donde θ es obtenida de φ reemplazando uniformemente variables proposicionales de φ por fórmulas arbitrarias.
- *Generalización*: si φ es teorema, $\Box\varphi$ es teorema.

Una fórmula es **K**-*demostrable* o un teorema de **K** si ocurre como último elemento de alguna **K**-prueba. Si este es el caso, escribimos: $\vdash_{\mathbf{K}} \varphi$.

Vale la pena hacer la aclaración de que la decisión de agregar *todas* las tautologías proposicionales como axiomas es innecesario. Podríamos haber elegido un conjunto pequeño de de tautologías capaces de generar el resto mediante reglas de inferencia, pero este refinamiento es de poco interés para nuestro propósito.

Veamos algunas propiedades del sistema axiomático **K**.

Modus Ponens. Respecto a *modus ponens* hay dos puntos importantes para destacar. En primer lugar, *modus ponens* preserva validez. Esto es, si $\models \varphi$ y $\models \varphi \to \psi$, entonces $\models \psi$. *Modus ponens* también preserva otras dos propiedades: *satisfacibilidad global* (*i.e.*, si $\mathcal{M} \models \varphi$ y $\mathcal{M} \models \varphi \to \psi$, entonces $\mathcal{M} \models \psi$) y *satisfacibilidad* (*i.e.* si $\mathcal{M}, w \models \varphi$ y $\mathcal{M}, w \models \varphi \to \psi$, entonces $\mathcal{M}, w \models \psi$). Esto significa que *modus ponens* es no sólo una regla correcta para razonar sobre frames (*i.e.*, intuitivamente, modelos sin valuación), sino que también lo es para razonar acerca de modelos, tanto global como localmente.

Substitución Uniforme. La regla de *substitución uniforme* refleja el hecho de que la validez se abstrae de los efectos de valuaciones particulares: al reemplazar uniformemente una variable proposicional dentro de una fórmula válida, seguimos obteniendo una fórmula válida. Notemos que si bien *substitución uniforme* preserva *validez*, esta regla *no* preserva *satisfacibilidad global* (*i.e.*, validez en un modelo) ni tampoco *satisfacibilidad.* Por ejemplo, q es claramente obtenible de p vía *substitución uniforme*, sin embargo del hecho de que p sea globalmente verdadera en un cierto modelo no se desprende que q también lo sea. En resumen: *substitución uniforme* es una herramienta únicamente para generar nuevas fórmulas universalmente válidas a partir de otras.

Generalización. Finalmente, la regla de *generalización* (también conocida como de *necesitación*) "modaliza" fórmulas demostrables agregando $\Box$s delante. Notemos que *generalización* preserva validez y *satisfacibilidad global*, sin embargo no preserva satisfacibilidad: por el sólo hecho de que p sea cierto en un estado, no podemos concluir que p es cierta en todo estado accesible.

Axioma K. Se puede ver fácilmente que el axioma K es válido. Si en un estado w de un modelo $\mathcal{M}$ todos los estados accesibles cumplen $p \to q$, esto equivale a decir $\mathcal{M}, w \models \Box(p \to q)$. Además, en el caso de que todos estos estados accesibles desde w cumplan también p esto significaría necesariamente que, dado que cumplen $p \to q$, también deben cumplir q, y esto determina que w satisface la fórmula $\Box p \to \Box q$. Hemos mostrado que cada vez que se cumple $\Box(p \to q)$ se debe cumplir $\Box p \to \Box q$, es decir $\mathcal{M}, w \models \Box(p \to q) \to (\Box p \to \Box q)$. Dado que no asumimos ninguna hipótesis, esta aseveración vale para w y $\mathcal{M}$ arbitrarios, y esto significa que la fórmula K es universalmente válida: $\models \Box(p \to q) \to (\Box p \to \Box q)$.

¿Cuál es el significado intuitivo de este axioma? K es también llamado el *axioma de distribución*, y es importante porque nos permite transformar una fórmula de tipo box $\Box(\varphi \to \psi)$ en una implicación $\Box\varphi \to \Box\psi$. Esta distributividad hace posible que haya más pasos de razonamiento puramente proposicional. Por ejemplo, supongamos que estamos tratando de probar $\Box\psi$, y hemos logrado construir una secuencia de prueba que contiene a $\Box(\varphi \to \psi)$ y $\Box\varphi$. Si pudiéramos aplicar la regla de *modus ponens* debajo del scope del $\Box$, podríamos probar $\Box\psi$. Esto es exactamente lo que la distributividad de K nos permite hacer: a partir de K, vía *substitución uniforme*, podemos probar $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$. Pero entonces la aplicación de *modus ponens* prueba $\Box\varphi \to \Box\psi$, y una segunda aplicación prueba $\Box\psi$ como queríamos.

El axioma K resulta intuitivo (como se esperaría) interpretando $\Box$ en términos de *necesidad*. Si es necesario que cada vez que ocurra p también deba ocurrir q, entonces si además se da el caso de que es necesario p, esto significará que será necesario q. Por ejemplo, si es necesario que haya electricidad cada vez que enciendo la lámpara, y es necesario encender la lámpara, entonces es necesario que haya electricidad: $\Box(\mathsf{lámpara} \to \mathsf{electricidad}) \to (\Box\,\mathsf{lámpara} \to \Box\,\mathsf{electricidad})$.

Podemos pensar que mientras K nos permite utilizar razonamiento clásico dentro de contextos modales, la regla de *necesitación* crea nuevos contextos para trabajar. Las pruebas modales surgen de la interacción entre estos dos mecanismos.

1.4 Bisimulaciones

Tiene sentido ahora preguntarse cuán expresivo es el lenguaje modal básico recién presentado. El poder expresivo de un lenguaje es por lo general medido en términos de las distinciones que puede hacer. Un lenguaje con solamente dos expresiones "gustar" y "disgustar", podría proveer una clasificación del mundo muy simplificada, mientras que un lenguaje más rico en aprobación y desaprobación podría hacer posible extraer distinciones más sutiles dentro de las situaciones aceptadas y rechazadas[3]. Entonces, ¿qué distinciones pueden hacer los lenguajes modales?

Un lenguaje modal (de hecho *cualquier* lenguaje lógico) puede distinguir entre algunos pares de modelos, pero no entre *todos*.

Por ejemplo, tomemos los modelos de la figura 1.1 y comparemos el punto s de $\mathcal{M}$ y el punto u de $\mathcal{K}$. Como cada uno de estos estados tiene sentido sólo en su modelo asociado, los notaremos $(\mathcal{M}, s)$ y $(\mathcal{K}, u)$ respectivamente[4]. A simple vista estas dos figuras parecen representar modelos muy disímiles. Pero ¿es posible distinguir *modalmente* $(\mathcal{M}, s)$ y $(\mathcal{K}, u)$, con la valuación vacía? Es decir, ¿es posible encontrar una fórmula φ (del lenguaje modal

[3] Consideremos, por ejemplo, la expresión coloquial "pse. . . ".
[4] Esta notación se conoce como *pointed models*.

básico) que sea verdadera en el punto s (*i.e.*, $\mathcal{M}, s \models \varphi$) pero falsa en el punto u (*i.e.*, $\mathcal{K}, u \not\models \varphi$)?

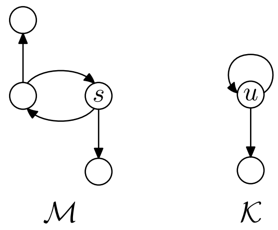

Fig. 1.1: $(\mathcal{M}, s)$ y $(\mathcal{K}, u)$ no son modalmente distinguibles

Notemos que es fácil distinguir estos modelos si podemos usar lógica de primer orden: todos los puntos en $\mathcal{M}$ (incluyendo s) son irreflexivos, mientras que el punto u de $\mathcal{K}$ es reflexivo, por lo tanto la fórmula de primer orden Rxx no es satisfacible (bajo ninguna asignación de variables posible) en el modelo $\mathcal{M}$, pero es satisfecha en $\mathcal{K}$ cuando u es asignada a x. Sin embargo, no vamos a encontrar una fórmula en el lenguaje modal básico que distinga estos modelos en esos puntos. Más adelante veremos cómo podemos probarlo.

Comparemos ahora nuestro modelo $\mathcal{M}$ con el modelo $\mathcal{N}$ de la figura 1.2. ¿Es posible distinguir modalmente $(\mathcal{M}, s)$ y $(\mathcal{N}, t)$?

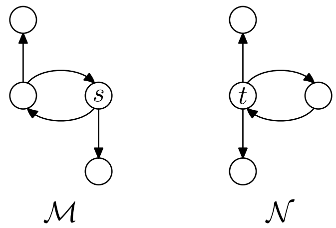

Fig. 1.2: $(\mathcal{M}, s)$ y $(\mathcal{N}, t)$ son modalmente distinguibles

Observemos que hay estados que no están relacionados con ningún otro estado tanto en $\mathcal{M}$ como en $\mathcal{N}$ (llamémoslos *estados terminales* o *estados deadlock*). Estos estados están caracterizados modalmente por la fórmula $\Box\bot$: como $\bot$ es falso en todo estado del modelo, $\Box\bot$ será satisfecho (por vacuidad) sólo por aquellos mundos no relacionados con ningún otro. Ahora consideremos que si tomamos cualquier relación desde $(\mathcal{M}, s)$ podemos llegar en un paso a un estado terminal ($\Box\bot$) o bien a un estado que en un paso puede llegar a un estado terminal ($\Diamond\Box\bot$). La fórmula $\Box(\Box\bot \vee \Diamond\Box\bot)$ expresa esta noción ya que dice que todo sucesor alcanza un estado terminal en 0 ó 1 paso. Pero esto último no es cierto para $(\mathcal{N}, t)$, si realizamos una transición hacia la derecha, llegaremos a un estado que no es terminal, y en donde necesitamos 2 pasos como mínimo para llegar a un estado terminal.

¿Cuándo son dos modelos modalmente indistinguibles? Como vimos en los ejemplos anteriores, esta no es una pregunta trivial. Existen modelos a primera vista diferentes ($(\mathcal{M}, s)$ y $(\mathcal{K}, u)$) que son modalmente equivalentes mientras que otros, quizás intuitivamente más similares ($(\mathcal{M}, s)$ y $(\mathcal{N}, t)$) que pueden distinguirse modalmente. Además esa pregunta es importante también prácticamente, por ejemplo interpretándolo desde el punto de vista de procesos, se traduce en ¿cómo reconocer dos diagramas de transición que son del mismo proceso? Los modelos $\mathcal{M}$ y $\mathcal{K}$ de la figura 1.1 parecen representar el mismo proceso cuando observamos las posibles acciones y deadlocks. Sin embargo, $\mathcal{M}$ y $\mathcal{N}$ en la figura 1.2 son diferentes, como dijimos, no todo estado en $\mathcal{N}$ puede acceder a un deadlock inmediato.

La relación que caracteriza la equivalencia semántica para la lógica de primer orden es la de isomorfismo. Dado que la lógica modal es un fragmento de la lógica de primer orden, el isomorfismo preserva la equivalencia modal, pero es una relación demasiado fuerte. En la figura 1.1 ambos modelos son claramente no-isomorfos, y sin embargo son modalmente equivalentes.

Por supuesto, no hay una respuesta única a la pregunta de cuándo dos estructuras son la misma. Primero debemos estipular alguna noción de equivalencia entre modelos que sea apropiada para lenguajes modales. Este es el propósito de la siguiente definición (formulada por primera vez por van Benthem) de *bisimulación*.

Definición 1.6 (Bisimulación)**.** Una *bisimulación* entre modelos $\mathcal{M} = \langle W, R, P\rangle$ y $\mathcal{M}' = \langle W', R', P'\rangle$ es una relación binaria no vacía E entre sus dominios ($E \subseteq W \times W'$) tal que si wEw' entonces:

Armonía atómica: w y w' satisfacen los mismos símbolos proposicionales,

Zig: si Rwv, entonces existe un punto v' en $\mathcal{M}'$ tal que vEv' y $R'w'v'$, y

Zag: si $R'w'v'$, entonces existe un punto v en $\mathcal{M}$ tal que vEv' y Rwv.

Diremos que dos modelos $\mathcal{M}$ y $\mathcal{N}$ son bisimilares si hay una bisimulación entre ellos, y diremos que dos estados w y v son bisimilares si están relacionados por alguna bisimulación entre sus modelos, y lo notaremos de la siguiente forma: $\mathcal{M}, w \leftrightarroweq \mathcal{N}, v$. Cuando $\mathcal{M}$ y $\mathcal{N}$ son obvios en el contexto notaremos directamente $w \leftrightarroweq v$.

La definición 1.6 nos dice que dos estados son bisimilares si hacen verdadera la misma información atómica, y sus posibilidades de transición pueden aparearse.

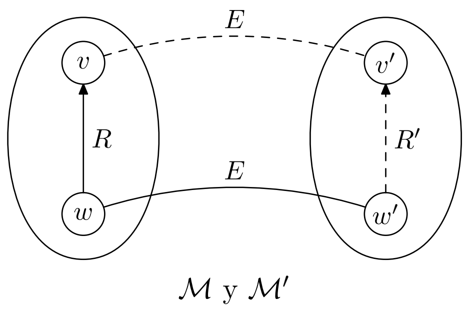

Fig. 1.3: La condición *zig*.

Exploremos un poco más la definición 1.6. La figura 1.3 muestra el contenido de la condición *zig*. Supongamos que conocemos que wEw' y Rwv. Entonces la condición *zig* especifica que es posible encontrar un estado v' que "completa el cuadrado". Notar la simetría entre las condiciones *zig* y *zag*.

Volviendo a los modelos $\mathcal{M}$, $\mathcal{N}$ y $\mathcal{K}$ considerados antes, se puede ver fácilmente que $\mathcal{M}$ y $\mathcal{K}$ son bisimilares: la línea punteada de la figura 1.4 muestra la bisimulación necesaria

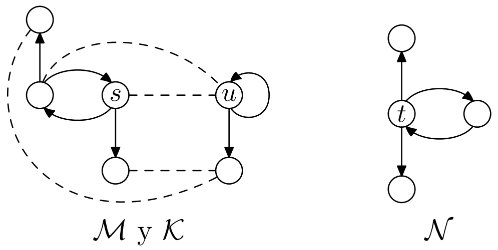

Fig. 1.4: $(\mathcal{M}, s)$ y $(\mathcal{K}, u)$ son bisimilares, y $(\mathcal{K}, u)$ y $(\mathcal{N}, t)$ no lo son.

que relaciona los dos puntos en cuestión. Sin embargo, no existe bisimulación Z entre $\mathcal{M}$ y $\mathcal{N}$ tal que $(s, t) \in Z$. Si bien la información atómica en $(\mathcal{M}, s)$ es la misma que en $(\mathcal{N}, t)$, comprobemos que la condición *zag* de la definición 1.6 no se cumple. Tomemos la transición hacia la derecha de t (llamémoslo $t_\rightarrow$), entonces el estado debajo de s (llamémoslo $s_\downarrow$) o bien el estado a la izquierda de s ($s_\leftarrow$) deben estar en la relación de bisimulación. Notemos que $(s_\downarrow, t_\rightarrow)$ no pueden estar en Z porque no se cumpliría la relación *zag*: desde $t_\rightarrow$ puedo acceder a un estado en donde vale $\top$ pero no puedo hacer lo mismo desde $s_\downarrow$. Entonces debe ser que $(s_\leftarrow, t_\rightarrow) \in Z$. En este caso, podríamos aplicar la condición *zig* eligiendo el estado hacia arriba de $s_\leftarrow$ ($s_\uparrow$) y el único estado accesible desde $t_\rightarrow$ que es t, y debe suceder que $(s_\uparrow, t) \in Z$. Pero nuevamente estamos en un estado terminal en $\mathcal{M}$ y en un no-terminal en $\mathcal{N}$: podemos acceder a un estado en donde vale $\top$ desde t pero no desde $s_\uparrow$ con o cual no se cumplirá la condición *zag*. Demostramos así que no existe una bisimulación entre $(\mathcal{M}, s)$ y $(\mathcal{N}, t)$.

El siguiente teorema formaliza las intuiciones detrás de la noción de bisimulación: las fórmulas modales son invariantes bajo bisimulación.

Teorema 1.1. [2, Teorema 2.20] Sean $\mathcal{M} = \langle W, R, P\rangle$ y $\mathcal{M}' = \langle W', R', P'\rangle$ modelos de Kripke, y $\mathcal{M}, w \leftrightarroweq \mathcal{M}', w'$ para $w \in W$ y $w' \in W'$, entonces w y w' satisfacen las mismas fórmulas modales básicas en la semántica de Kripke.

Demostración. Por inducción en la complejidad de φ. El caso en donde φ es una variable proposicional se sigue de la cláusula de *armonía atómica* de la definición 1.6 de bisimulación. Los casos booleanos son inmediatos por hipótesis inductiva.

Para las fórmulas de la forma $\Diamond\psi$, tenemos que $\mathcal{M}, w \models \Diamond\psi$ sii existe un v en $\mathcal{M}$ tal que Rwv y $\mathcal{M}, v \models \psi$. Como $w \leftrightarroweq w'$ tenemos por la cláusula de *zig* de la definición 1.6 que existe un v' en $\mathcal{M}'$ tal que $R'w'v'$ y $v \leftrightarroweq v'$. Por la hipótesis inductiva, $\mathcal{M}', w' \models \psi$, y por lo tanto $\mathcal{M}', w' \models \Diamond\psi$. Para la otra dirección usamos la cláusula de *zag* de la definición 1.6. □

Este resultado es esencial para asegurar que la noción de bisimulación es correcta. En otras palabras, el teorema nos asegura que la equivalencia estructural de la definición 1.6 captura efectivamente la equivalencia semántica de la lógica modal básica.

Si bien el teorema 1.1 prueba que de alguna forma las clases de equivalencia definidas por la bisimulación están incluidas o son más granulares que la relación de equivalencia semántica, nada sabemos aún del recíproco. ¿Podría ocurrir que dos estados modalmente equivalentes no sean bisimilares? La respuesta es *no* en la mayoría de los casos de interés:

Definición 1.7 (Imagen finita)**.** Diremos que un modelo $\mathcal{M} = \langle W, R, V\rangle$ tiene imagen finita si su relación de accesibilidad es de imagen finita:

$$\text{para todo } x \in W \text{ se cumple } \#\{w \mid Rxw\} < \infty.$$

Teorema 1.2 (Teorema de Hennessy-Milner)**.** [2, Teorema 2.24] Sean $\mathcal{M} = \langle W, R, V\rangle$ y $\mathcal{M}' = \langle W', R', V'\rangle$ dos modelos con imagen finita. Entonces, para cada $w \in W$ y $w' \in W'$: $\mathcal{M}, w \leftrightarroweq \mathcal{M}', w'$ sii w y w' son modalmente equivalentes.

Demostración. La dirección de izquierda a derecha se sigue del teorema 1.1; para la otra dirección, probaremos que la relación de equivalencia modal entre fórmulas (llamémosla E) es en sí misma una relación de bisimulación.

$$Eww' \quad \text{sii} \quad \text{para toda } \varphi : \mathcal{M}, w \models \varphi \text{ sii } \mathcal{M}', w' \models \varphi$$

Por hipótesis y nuestra definición de E tenemos que E es una relación no vacía en $W \times W'$ tal que $(w, w') \in E$.

La primera condición de bisimulación (armonía atómica) es inmediata. Para la segunda (*zig*), asumamos que w y w' son modalmente equivalentes y Rwv. Trataremos de llegar a una contradicción asumiendo que no existe v' en $\mathcal{M}'$ con $R'w'v'$ donde v y v' son modalmente equivalentes. Sea $S' = \{u' \mid R'w'u'\}$. Notemos que S' debe ser no-vacío, de otra forma $\mathcal{M}', w' \models \Box\bot$, que contradiría al hecho de que w y w' son modalmente equivalentes dado que $\mathcal{M}, w \models \Diamond\top$. Ahora bien, como $\mathcal{M}'$ es de imagen finita, S' debe ser finito, por ejemplo $S' = \{w'_1, \ldots, w'_n\}$. Por nuestra asunción, para cada $w'_i \in S'$ debe existir una fórmula ψ_i tal que $\mathcal{M}, v \models \psi_i$ pero $\mathcal{M}', w'_i \not\models \psi_i$. Entonces,

$$\mathcal{M}, w \models \Diamond(\psi_1 \wedge \ldots \wedge \psi_n) \quad \text{y} \quad \mathcal{M}', w' \not\models \Diamond(\psi_1 \wedge \ldots \wedge \psi_n),$$

que contradice nuestra asunción de que w y w' son modalmente equivalentes.

La tercera condición de la definición 1.6 (*zag*) puede ser probada de forma similar. □

Notemos en primer lugar cómo la demostración hace uso de la hipótesis de que $\mathcal{M}$ es un modelo de imagen finita. De otra forma, no hubiéramos podido construir una *fórmula* de tamaño finito para llegar a la contradicción. Resulta ser que el teorema *no* es válido para modelos de imagen infinita.

Teorema 1.3. El teorema 1.2 no es válido para la clase de todos los modelos.

Demostración. Sean $\mathcal{M} = \langle W, R, V\rangle$ y $\mathcal{N} = \langle W', R', V'\rangle$ como en la figura 1.5, donde ambos modelos tienen forma de estrella y cada flecha denota una R-transición. Tanto V como V' son valuaciones vacías (*i.e.*, funciones constantes vacías). Por cada número natural n, tanto $\mathcal{M}$ como $\mathcal{N}$ tienen una rama de longitud n; la diferencia entre estos dos modelos es que $\mathcal{N}$ además tiene una rama adicional de longitud infinita.

Se puede ver que dada una fórmula φ, se cumple que $\mathcal{M}, w \models \varphi$ sii $\mathcal{N}, w' \models \varphi$. Para esto hay que ver que si φ tiene k operadores modales, se pueden considerar solamente los caminos de longitud menor o igual a k e ignorar el resto. Cada vez que restringimos nuestros modelos de esta forma, independientemente de k, obtenemos modelos equivalentes. Por esto es que $(\mathcal{M}, w)$ y $(\mathcal{N}, w')$ resultan modalmente equivalentes.

Sin embargo estos dos modelos no son bisimilares. Para ver esto, supongamos que existe una bisimulación Z y derivemos una contradicción de este hecho.

Como wZw', debe haber un sucesor de w, sea v_0, que está relacionado al primer estado v'_0 de la rama infinita que parte de w'. Supongamos que el camino (maximal) determinado por los primeros dos estados w y v_0 es de longitud n, y sea $w, v_0, \ldots, v_{n-1}$ los puntos sucesivos de este camino. Usando la condición de *zig* de la definición 1.6 de bisimulación encontraremos

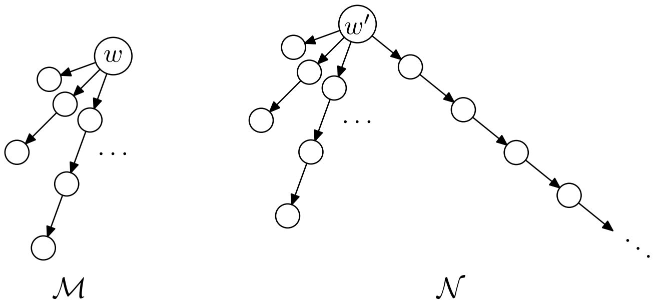

Fig. 1.5: Equivalentes pero no bisimilares.

los puntos $v'_1, \ldots, v'_{n-1}$ en el camino infinito que parte de w', tales que $v'_0 R' v'_1 R' \ldots R' v'_{n-1}$ y $v_i Z v'_i$ para cada i. Ahora bien, v'_{n-1} tiene un sucesor, pero v_{n-1} no lo tiene, y por lo tanto fallará la condición *zag* de bisimulación. Por lo tanto, $(\mathcal{M}, w) \not\simeq (\mathcal{N}, w')$. $\square$

Si agregáramos a nuestro lenguaje la posibilidad de tener conjunciones *infinitas* (*i.e.*, permitirnos fórmulas de la forma $\bigwedge\{\varphi_i \mid i \in I\}$ en nuestro lenguaje, con $\#I \leq \infty$), entonces podemos mostrar que dos modelos son bisimilares si y sólo si son modalmente equivalentes. Es decir que no haría falta la hipótesis de modelos de imagen finita. Para probarlo, podemos usar la demostración del teorema 1.2 de Hennessy-Milner pero en vez de construir una conjunción *finita* de las fórmulas que no se satisfacen en los *finitos* sucesores de w', construiremos una conjunción posiblemente *infinita* de las posiblemente infinitas fórmulas que no se satisfacen en cada uno de los posiblemente *infinitos* sucesores de w'. Como el resto de la prueba no utiliza la hipótesis de imagen finita, permanecerá sin cambios.

La definición 1.6 de bisimulación es una buena caracterización de la invarianza semántica entre los modelos de la lógica modal básica: si bien existen situaciones en las que no caracteriza la equivalencia lógica, esto se debe al carácter *finito* de las fórmulas del lenguaje. Pero ¿servirá también para otros lenguajes modales? Veremos que en las extensiones de la lógica modal básica que hemos visto, la idea de bisimulación consistirá en una extensión de la definición 1.6 vista.

Operadores temporales Hemos visto dos formas de definir a los operadores temporales de Prior. La primera de ellas, utilizando una única relación de accesibilidad que es la que nos da la noción de orden temporal de mundos posibles. El comportamiento de los operadores temporales se vio reflejado en la nuevas definiciones de satisfacibilidad que agregamos. Notemos que si bien F se comporta exactamente como un $\Diamond$ de la lógica modal básica, P navega la relación de accesibilidad en sentido inverso. Las condiciones *zig* y *zag* de la definición 1.6 de bisimulación aparean los mundos moviéndose únicamente en el sentido de la relación de accesibilidad. Entonces deberemos agregar dos condiciones similares a *zig* y *zag* teniendo en cuenta la navegabilidad de R en forma inversa para poder simular el operador P.

Definición 1.8 (Bisimulación para lógica temporal)**.** Definiremos la bisimulación entre modelos $\mathcal{M} = \langle W, R, P\rangle$ y $\mathcal{M}' = \langle W', R', P'\rangle$ para la lógica temporal como una relación binaria no vacía E entre sus dominios ($E \subseteq W \times W'$) tal que si wEw' entonces:

Armonía atómica: w y w' satisfacen los mismos símbolos proposicionales,

Zig: si Rwv, entonces existe un punto v' en $\mathcal{M}'$ tal que vEv' y $R'w'v'$,

Zag: si $R'w'v'$, entonces existe un punto v en $\mathcal{M}$ tal que vEv' y Rwv,

Zig^{-1}: si Rvw, entonces existe un punto v' en $\mathcal{M}'$ tal que vEv' y $R'v'w'$, y

Zag^{-1}: si $R'v'w'$, entonces existe un punto v en $\mathcal{M}$ tal que vEv' y Rvw.

Por otro lado, hemos mostrado que la lógica temporal también se puede definir como una extensión de la lógica modal básica con dos operadores, restringiendo las relaciones asociadas a cada operador. En otras palabras, P y F son operadores modales independientes pero sólo consideramos modelos donde las relaciones de accesibilidad son una la inversa de otra. Sobre esta clase de modelos, la noción de bisimulación seguirá siendo la misma de la definición 1.6 (con las condiciones *zig* y *zag* sobre ambas relaciones de accesibilidad R_1 y R_2).

Veremos ahora el caso del operador universal.

Operador universal Intuitivamente la nueva definición partirá a los modelos en más clases de equivalencia, dado que hay estados de modelos que en la lógica modal básica eran bisimilares pero que no serán modalmente equivalentes haciendo uso de la modalidad universal. Pero, ¿qué restricción debemos agregar a la noción de bisimulación para hacer equivalentes a estos modelos?

Dados dos modelos $\mathcal{M}_1$ y $\mathcal{M}_2$, quisiéramos asegurarnos que si hay alguna propiedad φ que se cumple *universalmente* en un modelo (*i.e.*, en todos los estados de $\mathcal{M}_1$) entonces se cumple universalmente en el otro. Y sin embargo, no queremos perder la noción anterior de bisimulación para la lógica modal básica: notemos que dos modelos bisimilares bajo esta nueva noción deberían ser también bisimilares para la lógica modal básica.

Si supiéramos que existe alguna bisimulación (en el sentido de la lógica modal básica) que relaciona a *todos* los estados de $\mathcal{M}_1$ con estados de $\mathcal{M}_2$, nos aseguraríamos que si φ es universalmente cierta en $\mathcal{M}_1$, también se cumple universalmente en $\mathcal{M}_2$. Esto es justamente lo que vamos a hacer. Modificaremos nuestra relación de bisimulación añadiendo la condición de que la bisimulación debe ser *total* en ambos sentidos: cada elemento de W_1 debe ser bisimilar a algún elemento de W_2, y cada elemento de W_2 debe ser bisimilar a alguno de W_1.

Con estos ejemplos hemos visto cómo la relación de bisimulación se "ajusta" al lenguaje con el que estamos tratando. En los próximos capítulos, definiremos y estudiaremos las bisimulaciones de lenguajes con el objeto de comprenderlos y tener una visión más clara de su expresividad.

1.4.1 Caracterización de van Benthem

Un resultado importante que añade valor a la definición 1.6 de bisimulación es el conocido como *caracterización de van Benthem*.

Se puede ver que la definición 1.3 de satisfacibilidad en la lógica modal básica puede escribirse en términos de la lógica de primer orden. Basta notar que cada cláusula de la definición cuantifica existencial o universalmente sobre elementos del dominio y que un modelo relacional de Kripke es esencialmente un modelo de primer orden. Esta traducción de fórmulas de la lógica modal básica a fórmulas de primer orden puede hacerse usando a lo sumo una variable libre, y recibe el nombre de *traducción estándar*. Esta traducción puede hacerse en un *lenguaje de correspondencia* sin igualdad con una variable libre en una signatura con relación binaria, relaciones unarias y sin símbolos de función ni constantes.

Dada una fórmula $\alpha(x)$ de primer orden, diremos que es *invariante bajo bisimulaciones* si para todos los modelos y estados tales que $(\mathcal{M}, w) \leftrightarroweq (\mathcal{N}, v)$, resulta que $\mathcal{M} \models \alpha(x)[w]$ sii $\mathcal{N} \models \alpha(x)[v]$. Es decir, que el modelo $\mathcal{M}$ satisface $\alpha(x)$ en cualquier interpretación de primer orden en donde $x \mapsto w$, si y sólo si el modelo $\mathcal{N}$ satisface $\alpha(x)$ en toda interpretación donde $x \mapsto v$.

Finalmente, el teorema de caracterización de van Benthem establece que las fórmulas del lenguaje de primer orden del *lenguaje de correspondencia* son invariantes bajo bisimulación si y sólo si son equivalentes a la traducción estándar de una fórmula modal. Dicho de otra forma, si los modelos de primer orden preservan equivalencia de la fórmula $\alpha(x)$ al ser transformados mediante bisimulaciones, entonces necesariamente debe ser el caso de que $\alpha(x)$ es en realidad equivalente a la traducción estándar de φ, donde φ es una fórmula modal básica; y viceversa.

Este resultado otorga un gran valor teórico a la bisimulación, dado que afirma que la relación de bisimulación caracteriza con exactitud a las fórmulas de primer orden expresables en lógica modal. Es decir que el teorema de caracterización expresa cuál es *precisamente* el fragmento de primer orden que se corresponde con la lógica modal, en términos de una propiedad *estructural* de los modelos: bisimulación.

2. INTRODUCCIÓN A NEIGHBOURHOOD SEMANTICS

Si bien la forma más usual de pensar lógicas modales es dentro de una semántica relacional de Kripke, existen otras alternativas de semánticas interesantes, algunas de ellas más débiles, como ser *Algebraic Semantics*, *Topological Semantics*, o *Neighbourhood Semantics*.

2.1 Motivación

Para muchas aplicaciones, la semántica relacional que presentamos en el capítulo 1 resulta demasiado fuerte, y es necesario utilizar una semántica más débil o general en donde se pueda tener mayor control del poder de razonamiento que se quiere otorgar al agente involucrado. Veamos en detalle un ejemplo de este problema.

2.1.1 Problema de la omnisciencia lógica

La lógica epistémica es la lógica del *conocimiento* y de la *creencia*. El lenguaje clásico de la lógica modal contiene los operadores de la lógica proposicional, junto con operadores de la forma "$\Box$". La interpretación epistémica que se espera de una sentencia de la forma $\Box\varphi$ es que el agente *conoce* o *cree* que φ. Sin embargo, los términos "agente", "conoce" y "cree" deben ser entendidos en un sentido muy laxo. Un "agente" puede ser cualquier tipo de objeto para el cual tiene sentido decir que posee información: humanos, pero también robots, bases de datos, y, en un sentido más abstracto, agentes de software o incluso aparatos electrónicos. Decir que un agente humano conoce y cree ciertas cosas no es muy controversial; atribuir conocimientos y creencias a agentes abstractos sí lo es. En este trabajo utilizaremos los términos "conocimiento", "creencia" e "información" de forma más o menos intercambiable. Es quizás conveniente leer el verbo "creer" como "tener la información de que" y "conocer" como una abreviatura de "tener la información *cierta* de que".

La semántica de los mundos posibles de Kripke da lugar a una noción de *conocimiento* y *creencia* bastante distorsionada. En particular, bajo esta semántica vale que si ψ es una consecuencia lógica de φ, entonces en todos los modelos en donde $\Box\varphi$ es verdadero (*i.e.*, donde el agente conoce que φ es verdadero), $\Box\psi$ es asimismo verdadero. Es decir, cumplen el axioma K[1]:

$$\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$$

Más aún, cumplen esquema de inferencia de *Generalización*[2]:

$$\text{si } \models \varphi, \text{ entonces } \models \Box\varphi$$

En otras palabras, la *creencia* (y el *conocimiento*) de un agente es cerrado bajo consecuencia lógica. Pero esto no se condice con la situación real: las personas no siempre ven las

[1] Ya hemos visto este axioma y comprobado su validez en la pág. 7 (sección 1.3).
[2] También visto en la sección 1.3.

consecuencias de sus creencias, e incluso saben que otras personas son limitadas en el mismo sentido [5]. Este hecho es particularmente evidente cuando echamos un vistazo a las verdades matemáticas: una vez que conocemos los axiomas básicos de la aritmética, no sabemos si el último teorema de Fermat es verdadero, a pesar de que es una consecuencia lógica de los axiomas. Esta desincronización ("mismatch") entre las semánticas y el concepto de conocimiento y creencia es conocido como *el problema de la omnisciencia lógica*.

Si las semánticas epistémicas son vistas como una propuesta de una teoría que describe el uso lingüístico de frases como "Juan sabe que ...", entonces el problema de la omnisciencia lógica claramente falsifica el análisis del conocimiento en la semántica de los mundos posibles de Kripke. "Pedro sabe que el teorema de Fermat es verdadero." es una oración que es una contingencia, y sin embargo la fórmula correspondiente es una tautología.

2.1.2 Otras aplicaciones

También para otras aplicaciones, la semántica de los mundos posibles de Kripke es demasiado fuerte, incluso considerando esquemas más débiles que el recién visto. Por ejemplo, en la semántica de los mundos posibles de Kripke $\Diamond(\varphi \vee \psi) \to (\Diamond\varphi \vee \Diamond\psi)$ es válido. Si desde un mundo w puedo alcanzar mediante la relación de accesibilidad un estado v en donde es cierto $\varphi \vee \psi$. Entonces si φ es cierto en v, $\Diamond\varphi$ se satisface en w; en caso contrario ψ es válido en v y $\Diamond\psi$ en w, por lo que $\Diamond\varphi \vee \Diamond\psi$ es cierto en w. Sin embargo si interpretamos $\Diamond\varphi$, en un contexto de teoría de juegos, como que el jugador tiene una estrategia que fuerza φ, entonces podríamos estar inclinados a rechazarlo: ¿Por qué el tener una estrategia para la disyunción implica tener una estrategia para uno de los disyuntos? Por ejemplo, supongamos que jugamos un juego con los siguientes movimientos: usted tiene permitido decidir entre ir al cine o ir a un concierto, y yo puedo decidir a qué película o concierto en particular ir. Entonces, yo puedo forzar $\mathsf{Amélie} \vee \mathsf{Bach}$, pero (como usted es quien determina la opción cine/concierto) yo no puedo determinar cuál de estas dos opciones ocurrirá efectivamente.

2.2 Definiciones básicas

Las consideraciones que discutimos en la sección anterior han llevado a buscar semánticas modales más débiles que la relacional. La más conocida de estas es *Neighbourhood Semantics*, en primera instancia introducida por Montague [8, 9] y Scott [10] en 1970, si bien años más tarde se introduciría una versión modificada. La idea principal de *Neighbourhood Semantics* es que cada punto w en un modelo está asociado a una colección de subconjuntos del dominio (estos son los *neighbourhoods* o vecindades de w) y una fórmula de la forma $\Box\varphi$ es verdadera en w si y sólo si el conjunto de los puntos en el modelo que satisfacen φ es una vecindad de w. Intuitivamente, esta parece ser la semántica más débil posible, ya que permitimos explícitamente que cada elemento del dominio defina qué fórmulas del tipo $\Box\varphi$ valida.

Formalicemos el tipo de estructuras con las que vamos a trabajar.

Definición 2.1 (Modelo minimal [3])**.** Decimos que $\mathcal{M} = \langle W, N, P\rangle$ es un *modelo minimal* cuando:

1. W es un conjunto no vacío.
2. N es una función de W a conjuntos de subconjuntos de W (*i.e.*, $N_w \subseteq \wp(W)$, para cada mundo w en W).

3. P es una función del conjunto de variables proposicionales a subconjuntos de W (*i.e.*, $P(p) \subseteq W$, para cada variable proposicional $p \in \mathsf{PROP}$).

Donde W es el dominio de mundos posibles, P es una función de asignación de valuación, y N una relación entre elementos del dominio y vecindades. Llamaremos N_w al conjunto de las vecindades del mundo w.

Históricamente han habido dos definiciones distintas de *Neighbourhood Semantics*, la primera de ellas introducida por Montague [8, 9] y Scott [10] en 1970. De aquí en más nos referiremos a esta noción como *Neighbourhood Semantics Strict*. La definicion alternativa fue introducida por Johan van Benthem en [1] y en escritos no publicados y es, hoy por hoy, la más conocida y habitual. Aquí la llamaremos *Neighbourhood Semantics Loose*.

De aquí en más nos referiremos al lenguaje *Neighbourhood Semantics Strict* con el símbolo $\mathcal{N}_=$ y al lenguaje *Neighbourhood Semantics Loose* con $\mathcal{N}_\subseteq$. De igual manera, y debido a que a lo largo del presente trabajo se va a hacer referencia tanto a $\mathcal{N}_=$ como a $\mathcal{N}_\subseteq$, introduciremos algunos símbolos para facilitar la lectura de los resultados. Cuando haya lugar a confusión o ambigüedades nos referiremos con $\models_=$ a la relación de satisfacibilidad de $\mathcal{N}_=$ y con $\models_\subseteq$ a la relación de satisfacibilidad de $\mathcal{N}_\subseteq$. Asimismo emplearemos el símbolo $⊟$ en las fórmulas que debemos interpretar bajo la semántica $\mathcal{N}_=$, y el símbolo $⧇$ para $\mathcal{N}_\subseteq$. Sin embargo el lector debe notar que el lenguaje subyacente de ambas lógicas es el mismo: el lenguaje modal básico.

Veamos en primera instancia la definición concreta de *Neighbourhood Semantics Strict*.

Definición 2.2. Definimos la relación de satisfacibilidad de $\mathcal{N}_=$ sobre un modelo minimal $\mathcal{M} = \langle W, N, P\rangle$ de la siguiente forma:

$$\mathcal{M}, w \models p \text{ sii } w \in P(p) \text{ y } p \in \mathsf{PROP}$$

$$\mathcal{M}, w \models \neg\varphi \;\text{ sii } \mathcal{M}, w \not\models \varphi$$

$$\mathcal{M}, w \models \varphi \vee \psi \;\text{ sii } \mathcal{M}, w \models \varphi \text{ ó } \mathcal{M}, w \models \psi$$

$$\mathcal{M}, w \models ⊟\varphi \;\text{ sii } \|\varphi\|^{\mathcal{M}} \in N_w.$$

Donde $\|\varphi\|^{\mathcal{M}}$ es la extensión de φ en $\mathcal{M}$ que dimos en la definición 1.4.

Como dijimos, la idea principal de esta semántica es que cada posible mundo w en un modelo minimal $\mathcal{M} = \langle W, N, P\rangle$ tiene asociado un conjunto de extensiones N_w que son en algún sentido *necesarias* o *conocidas* en w. Como podemos identificar una fórmula en $\mathcal{M}$ con el conjunto de mundos posibles de $\mathcal{M}$ que satisfacen esta fórmula, N_w se convierte en una colección de subconjuntos de W. Se debe tener en cuenta que N_w puede ser *cualquier* colección de fórmulas, incluyendo la colección vacía. La semántica no hace ninguna asunción sobre la naturaleza de N excepto que debe ser una función de W a $\wp(\wp(W))$.

Por otro lado, la definición de *Neighbourhood Semantics Loose* utiliza los mismos tipos de modelos minimales, pero difiere en la definición de satisfacibilidad para el operador modal.

Definición 2.3. Definimos la relación de satisfacibilidad de $\mathcal{N}_\subseteq$ sobre un modelo minimal $\mathcal{M} = \langle W, N, P\rangle$ de igual forma que para $\mathcal{N}_=$ salvo en el caso del operador modal:

$$\mathcal{M}, w \models ⧇\varphi \;\text{ sii } \exists X \in N_w \text{ tal que } X \subseteq \|\varphi\|^{\mathcal{M}}.$$

Esta semántica difiere de *Neighbourhood Semantics Strict* en la definición del operador box: en este caso alcanza simplemente con que cada elemento de una vecindad N_w cumpla φ para poder aseverar $\boxbox\varphi$.

Notemos que la interpretación del operador de necesidad en modelos minimales es simple y natural: diremos que una sentencia de la forma $\Box\varphi$ (obviemos por un instante las diferencias de notaciones entre $\boxminus$ y $\boxbox$) es verdadera en un mundo posible α en $\mathcal{M}$ en el caso de que la proposición expresada por φ –el conjunto $\|\varphi\|^{\mathcal{M}}$– está entre los elementos de N_α (en el caso de *Neighbourhood Semantics Strict*) o bien si está algún subconjunto de ésta (en el caso de *Neighbourhood Semantics Loose*).

Como se ve, la definición de *Neighbourhood Semantics Strict* (de aquí en más $\mathcal{N}_=$) es más fuerte que la de *Neighbourhood Semantics Loose* (de aquí en más $\mathcal{N}_\subseteq$), si $\boxminus p$ es satisfecho en un mundo w de un modelo minimal $\mathcal{M}$, entonces $\boxbox p$ también será satisfecha bajo *Neighbourhood Semantics Loose*[3].

2.3 Comparación con semántica de Kripke

Ambas nociones de *Neighbourhood Semantics* son una *generalización* de la semántica relacional. Para ver esto, notemos que dado cualquier modelo relacional $\mathcal{M} = \langle W, R, P\rangle$, podemos formar un modelo minimal $\mathcal{M}^N = \langle W, N^R, P\rangle$ con la relación de accesibilidad definida como $N^R_w = \{U \mid U \subseteq W \text{ y } Rw \subseteq U\}$, donde $Rw = \{v \mid Rwv\}$. Este modelo resulta equivalente al relacional tanto para $\mathcal{N}_=$ como $\mathcal{N}_\subseteq$ como veremos a continuación.

Lema 2.1. Dado una fórmula φ, un modelo minimal $\mathcal{M}$ y un mundo w en el modelo,

$$\mathcal{M}, w \models \varphi \quad \text{sii} \quad \mathcal{M}^N, w \models_= \varphi \quad \text{sii} \quad \mathcal{M}^N, w \models_\subseteq \varphi.$$

Demostración. Esto se puede demostrar fácilmente por inducción en la longitud de φ. El único caso de interés es el modal.

$\Rightarrow$) Si $\mathcal{M}, w \models \Box\varphi$, esto significa que $Rw \subseteq \|\varphi\|^{\mathcal{M}}$, y dado que en N^R_w están todos los superconjuntos de Rw, en particular tenemos a $\|\varphi\|^{\mathcal{M}}$, y por hipótesis inductiva $\|\varphi\|^{\mathcal{M}} = \|\varphi\|^{\mathcal{M}^N} \in N^R_w$, esto significa que $\mathcal{M}^N, w \models_= \Box\varphi$ y $\mathcal{M}^N, w \models_\subseteq \Box\varphi$.

$\Leftarrow$) Para la dirección contraria, asumamos primero que $\mathcal{M}^N, w \models_= \Box\varphi$, es decir $\|\varphi\|^{\mathcal{M}^N} \in N^R_w$. Por definición de N^R esto significa que $Rw \subseteq \|\varphi\|^{\mathcal{M}^N}$, aplicando la hipótesis inductiva, $Rw \subseteq \|\varphi\|^{\mathcal{M}}$, por lo que $\mathcal{M}, w \models \Box\varphi$.

Para el caso de $\mathcal{N}_\subseteq$ debemos notar que el hecho de que $\mathcal{M}^N, w \models_\subseteq \Box\varphi$ equivale a la existencia de $X \in N^R_w$ tal que $X \subseteq \|\varphi\|^{\mathcal{M}^N}$. Nuevamente por definición de N^R, Rw debe ser subconjunto de X, y por lo tanto también de $\|\varphi\|^{\mathcal{M}^N}$, lo que por hipótesis inductiva significa $Rw \subseteq \|\varphi\|^{\mathcal{M}}$, en otras palabras, $\mathcal{M}, w \models \Box\varphi$. □

Si bien, como mostramos, un modelo relacional se puede convertir en un modelo minimal equivalente, no podemos hacer la traducción opuesta. Consideremos el modelo $\mathcal{M} = \langle W, N, P\rangle$ tal que $W = \{t, u, v\}$, $P(p) = \{t\}$ y $P(q) = \{v\}$, y $N_u = \{P(p), PIMQ\}$, donde $PIMQ = \{u, v\}$ como se muestra en la figura 2.1. Notemos que $PIMQ$ es el conjunto de puntos donde $p \to q$ es verdadero. Por lo tanto, $\mathcal{M}, u \models \boxminus(p \to q)$, dado que $PIMQ \in N_u$. Además, $\mathcal{M}, u \models \boxminus p$, dado que $P(p) \in N_u$. Sin embargo $\mathcal{M}, u \not\models \boxminus q$, porque $P(q) \notin N_u$. Entonces,

[3] Esto *no* significa que cualquier fórmula satisfacible en *Neighbourhood Semantics Strict* es satisfacible en *Neighbourhood Semantics Loose*, como veremos más adelante.

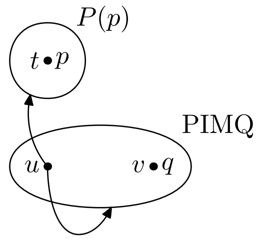

Fig. 2.1: Modelo $\mathcal{M}$ en donde no se cumple K.

$\mathcal{M}, u \not\models \boxed{=}(p \to q) \to (\boxed{=}p \to \boxed{=}q)$. Como esta fórmula es universalmente válida dentro de la semántica relacional (es una instancia del axioma K), no existe modelo relacional equivalente a $\mathcal{M}$. Es fácil verificar que este mismo modelo interpretado en $\mathcal{N}_{\subseteq}$ también cumple que $\mathcal{M}, u \not\models \boxed{\subseteq}(p \to q) \to (\boxed{\subseteq}p \to \boxed{\subseteq}q)$ y por lo tanto muestra que tampoco existe la traducción opuesta para $\mathcal{N}_{\subseteq}$.

Similarmente, el principio de inferencia de generalización característico de la semántica relacional (si $\models \varphi$ entonces $\models \Box\varphi$) no se cumple para $\mathcal{N}_{=}$ ni $\mathcal{N}_{\subseteq}$. Basta comprobar que en el modelo recién visto, $\mathcal{M}, t \not\models \Box\top$ para ninguna de las dos semánticas porque en particular $N_t = \emptyset$, y por lo tanto $\not\models \boxed{=}\top$ y $\not\models \boxed{\subseteq}\top$, mientras que $\models \top$ es cierto por definición.

De hecho, lo único que $\mathcal{N}_{=}$ preserva es una regla mucho más débil (monotonía sobre $\boxed{=}$): si $\models \varphi \leftrightarrow \psi$ entonces $\models \boxed{=}\varphi \leftrightarrow \boxed{=}\psi$. Esta regla es válida simplemente porque la relación de satisfacibilidad no es sintáctica, y la única manera de distinguir φ de ψ es sintácticamente. La semántica en particular se basa en las extensiones de las fórmulas, y siempre y cuando las extensiones de dos fórmulas sean las mismas, el modelo no podrá distinguir entre una y la otra. Notemos que $\mathcal{M}, w \models \boxed{=}\varphi \leftrightarrow \boxed{=}\psi$ es lo mismo que decir $\|\varphi\|^{\mathcal{M}} \in N_w$ sii $\|\psi\|^{\mathcal{M}} \in N_w$, pero dado que $\varphi \leftrightarrow \psi$, entonces $\|\varphi\|^{\mathcal{M}} = \|\psi\|^{\mathcal{M}}$. Además, por el mismo argumento, la regla de monotonía sobre $\boxed{\subseteq}$ también es válida para $\mathcal{N}_{\subseteq}$.

De esta forma, hemos visto que $\mathcal{N}_{=}$ y $\mathcal{N}_{\subseteq}$ no fuerzan omnisciencia lógica. En el capítulo 4 veremos que sin embargo $\mathcal{N}_{\subseteq}$ es una lógica más fuerte que $\mathcal{N}_{=}$.

En el siguiente capítulo iniciaremos nuestra investigación semántica de $\mathcal{N}_{=}$ y $\mathcal{N}_{\subseteq}$ en términos de bisimulaciones.

3. BISIMULACIONES

3.1 Neighbourhood Semantics Loose

En esta sección se tratarán las nociones de bisimulación para $\mathcal{N}_{\subseteq}$ y se mostrará la clase de modelos para la cual esta bisimulación coincide con la equivalencia modal entre modelos.

3.1.1 Bisimulación

La noción de bisimulación para $\mathcal{N}_{\subseteq}$ es intuitiva, sólo debemos encargarnos de aparear los neighbourhoods y recursivamente los elementos incluidos en estos.

Definición 3.1. Sean $\mathcal{M} = \langle W, N, P\rangle$, $\mathcal{M}' = \langle W', N', P'\rangle$ dos modelos minimales. Una *bisimulación minimal loose* entre estos dos modelos es una relación $Z \subseteq W \times W'$ no vacía tal que si xZx' entonces:

1. $\forall p \in \mathsf{PROP}$, $x \in P(p) \iff x' \in P'(p)$
2. (Zig) Para cada T en N_x debe existir T' en $N'_{x'}$ que verifique: $\forall w' \in T'\ \exists w \in T$ tal que wZw'
3. (Zag) Para cada T' en $N'_{x'}$ debe existir T en N_x que verifique: $\forall w \in T\ \exists w' \in T'$ tal que wZw'

Verifiquemos que esta bisimulación preserva equivalencia modal entre modelos para $\mathcal{N}_{\subseteq}$.

Proposición 3.1. Sean $\mathcal{M} = \langle W, N, P\rangle$, $\mathcal{M}' = \langle W', N', P'\rangle$ dos modelos minimales, Z una bisimulación minimal loose entre ellos, $w \in W$, $w' \in W'$ y wZw'. Entonces, para cada fórmula modal $\varphi \in \mathcal{N}_{\subseteq}$:

$$\mathcal{M}, w \models \varphi \iff \mathcal{M}', w' \models \varphi$$

Demostración. Por inducción en φ:

<u>Caso base</u>: Si φ es una variable proposicional, la condición 1 de bisimulación asegura que $w \in P(\varphi) \iff w' \in P'(\varphi)$, entonces $\mathcal{M}, w \models \varphi \iff \mathcal{M}', w' \models \varphi$.

<u>Paso inductivo 1</u>: Si φ es una fórmula booleana de la forma $\neg\psi$ ó $\psi_1 \wedge \psi_2$, sólo hace falta aplicar la hipótesis inductiva.

<u>Paso inductivo 2</u>: El caso modal: $\varphi = \boxdot\psi$.

$\Rightarrow$) Supongamos que $\mathcal{M}, w \models \boxdot\psi$, entonces $\exists X \in N_w$ tal que $X \subseteq \|\psi\|^{\mathcal{M}}$. Por la condición 2 de bisimulación (*zig*) debe haber un $X' \in N'_{w'}$ de forma tal que $\forall x' \in X'\ \exists x \in X$ tal que xZx'. Sea $x' \in X' \in N'_{w'}$ arbitrario; como cada $t \in X$ verifica $\mathcal{M}, t \models \psi$ y además x y x' son bisimilares, podemos aplicar la hipótesis inductiva: $\mathcal{M}', x' \models \psi$. Esto ocurre para todos los $x' \in X'$, entonces $X' \subseteq \|\psi\|^{\mathcal{M}'}$ lo que significa que $\mathcal{M}', w' \models \boxdot\psi$.

$\Leftarrow$) Ahora supongamos que $\mathcal{M}', w' \models \Box\psi$. Por la condición 3 de bisimulación (*zag*) debe existir $X \in N_w$ de forma tal que $\forall x \in X\ \exists x' \in X'$ tal que xZx'. Como en el caso $\Rightarrow$) aplicamos la hipótesis inductiva: $\mathcal{M}, x \models \psi$ para cada $x \in X$, entonces $X \subseteq \|\psi\|^{\mathcal{M}}$ lo que significa que $\mathcal{M}, w \models \Box\psi$. □

De igual forma que en la lógica modal relacional, es importante que esta definición de bisimulación tenga un resultado paralelo al teorema 1.2 de Hennessy-Milner. De esta forma comprobaremos cuán próxima es la definición 3.1 de bisimulación con la noción de equivalencia modal, y cuáles son las condiciones necesarias exactas para que ambas coincidan.

Definición 3.2. Un modelo minimal $\mathcal{M} = \langle W, N, P\rangle$ tiene la propiedad de:

$$\begin{array}{rll} \textit{neighbourhood-finito} & \text{sii} & \forall w \in W\ \forall T \in N_w : |T| < \infty, \\ \textit{co-neighbourhood-finito} & \text{sii} & \forall w \in W\ \forall T \in N_w : |W \setminus T| < \infty, \text{ e} \\ \textit{imagen-finita} & \text{sii} & \forall w \in W : |N_w| < \infty. \end{array}$$

Proposición 3.2. Sean $\mathcal{M} = \langle W, N, P\rangle$, $\mathcal{M}' = \langle W', N', P'\rangle$ dos modelos minimales con las propiedades de imagen-finita y neighbourhood-finito. Sean $w \in W$ y $w' \in W'$ tales que para todo $\varphi \in \mathcal{N}_{\subseteq} : \mathcal{M}, w \models \varphi$ sii $\mathcal{M}', w' \models \varphi$ (*i.e.*, w y w' son modalmente equivalentes). Entonces w y w' son bisimilares (*i.e.*, $\exists Z$ bisimulación minimal loose entre $\mathcal{M}$ y $\mathcal{M}'$ tal que wZw').

Demostración. Proponemos que la definición que necesitamos está definida por:

$$xZx' \text{ sii } \forall\varphi : \mathcal{M}, x \models \varphi \iff \mathcal{M}', x' \models \varphi.$$

Debemos probar que Z es efectivamente una bisimulación. Por definición, $(w, w') \in Z$ lo que asegura que Z es no vacía. Debemos ahora chequear las condiciones 1 a 3 de la definición 3.1.

1. La primera condición de bisimulación está satisfecha por definición de Z.

2. Para probar que la segunda condición es satisfecha, supongamos *ad absurdum* que no lo es, por lo cual w y w' son modalmente equivalentes y en relación wZw', $T \in N_w$ pero $\forall T' \in N'_{w'}$ la condición 2 falla. Sea $T = \{h_1, \ldots, h_n\}$ y $N'_{w'} = \{T'_1, \ldots, T'_m\}$.

   Como la condición debe ser falsificada, para cualquier $T'_i \in N'_{w'}$ debe haber un elemento $h'_i \in T'_i$ tal que no es bisimilar con ningún $h_j \in T$. Esto significa que debe existir una fórmula $\psi_{i,j}$ tal que $\mathcal{M}, h_j \models \psi_{i,j}$ pero $\mathcal{M}', h'_i \not\models \psi_{i,j}$.

   Sea $\varphi = \bigwedge_{i=1}^{m} \bigvee_{j=1}^{n} \psi_{i,j}$. Claramente $T \subseteq \|\varphi\|^{\mathcal{M}}$. Pero para cada T'_k hay un elemento $h'_k \notin \|\varphi\|^{\mathcal{M}'}$, porque en particular $\mathcal{M}', h'_k \not\models \bigvee_{j=1}^{n} \psi_{k,j}$.

   Por lo tanto, $\mathcal{M}, w \models \Box\varphi$ pero $\mathcal{M}', w' \not\models \Box\varphi$. Absurdo.

3. La tercera condición de bisimulación se prueba de manera análoga a la segunda condición.

□

La proposición 3.2 muestra que las condiciones de *neighbourhood-finito* e *imagen-finita* son suficientes para asegurar que las nociones de bisimilaridad y equivalencia modal coincidan. Veamos ahora que estas condiciones son en realidad necesarias.

Lema 3.1. La proposición 3.2 no es válida para modelos minimales sin la propiedad de neighbourhood-finito.

Demostración. Sea $\mathcal{M} = \langle W, N, P \rangle$, donde $W = \{w_i \mid i \geq 0\} \cup \{h\}$, $N_{w_0} = \{\{w_i \mid i > 0\} \cup \{h\}\}$, $N'_a = \emptyset$ para $a \neq w'_0$, y $\forall i > 0 : P(p_i) = \{w_1, \ldots, w_i, h\}$.

Y sea $\mathcal{M}' = \langle W', N', P' \rangle$, donde $W' = \{w'_i \mid i \geq 0\}$, $N'_{w'_0} = \{\{w'_i \mid i > 0\}\}$, $N'_a = \emptyset$ para $a \neq w'_0$, y $\forall i > 0 : P'(p_i) = \{w'_1, \ldots, w'_i\}$, como en la figura 3.1.

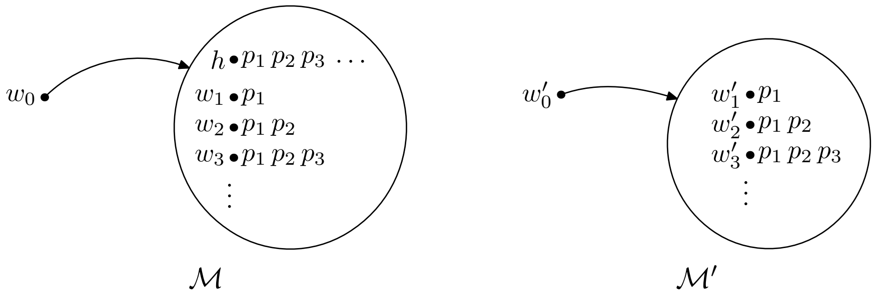

Fig. 3.1: $(\mathcal{M}, w_0)$ y $(\mathcal{M}', w'_0)$ son modalmente equivalentes pero no bisimilares.

Veamos que $(\mathcal{M}, w_0)$ y $(\mathcal{M}', w'_0)$ no son bisimilares, es decir, no existe una bisimulación Z tal que $w_0 Z w'_0$. Para esto, es fácil ver que h no es bisimilar a *ninguno* de los elementos en $N_{w'_0}$, esto implica que la condición 3 de la definición de bisimulación no es satisfecha: Dado el único neighbourhood de w'_0 (llamémoslo T'), no podemos encontrar un neighbourhood T de w_0 tal que todo elemento sea bisimilar a alguno de T'. El único neighbourhood de w'_0 no sirve, porque ninguno de sus elementos cumple la condición 1 de bisimulación (armonía atómica) con respecto a h.

Sin embargo, no hay una fórmula φ tal que $\mathcal{M}, w_0 \models \varphi$ y $\mathcal{M}', w'_0 \not\models \varphi$. Esto se debe a que φ es una fórmula finita. Si p_k es la proposición con k más grande en φ, entonces el mundo h de $\mathcal{M}$ no podrá ser diferenciado de w_k por esta fórmula.

Así w_0 y w'_0 son modalmente equivalentes, pero no existe ninguna bisimulación que los relacione. $\square$

Lema 3.2. La proposición 3.2 no es válida para modelos minimales sin la propiedad de imagen-finita.

Demostración. Sea $\mathcal{M} = \langle W, N, P \rangle$, donde $W = \{w_i \mid i \geq 0\} \cup \{h\}$, $N_{w_0} = \{\{w_i\} \mid i > 0\} \cup \{\{h\}\}$, $\forall i > 0 : P(p_i) = \{w_1, \ldots, w_i, h\}$.

Y sea $\mathcal{M}' = \langle W', N', P' \rangle$, donde $W' = \{w'_i \mid i \geq 0\}$, $N'_{w'_0} = \{\{w'_i\} \mid i > 0\}$, $\forall i > 0 : P(p_i) = \{w'_1, \ldots, w'_i\}$ como lo muestra la figura 3.2.

Notar que ambos modelos son de neighbourhood-finito. Además, w_0 y w'_0 son modalmente equivalentes: dado una fórmula φ y p_k la variable proposicional con mayor k que ocurre en φ, esta fórmula no podrá discriminar el neighbourhood que contiene h del neighbourhood que contiene w_k, y por lo tanto no podrá distinguir los dos modelos.

Sin embargo, los modelos no son bisimilares, dado que el neighbourhood singleton $\{h\} \in N_{w_0}$ no tiene otro neighbourhood correspondiente en $N'_{w'_0}$ donde la condición 2 de bisimulación sea satisfecha. $\square$

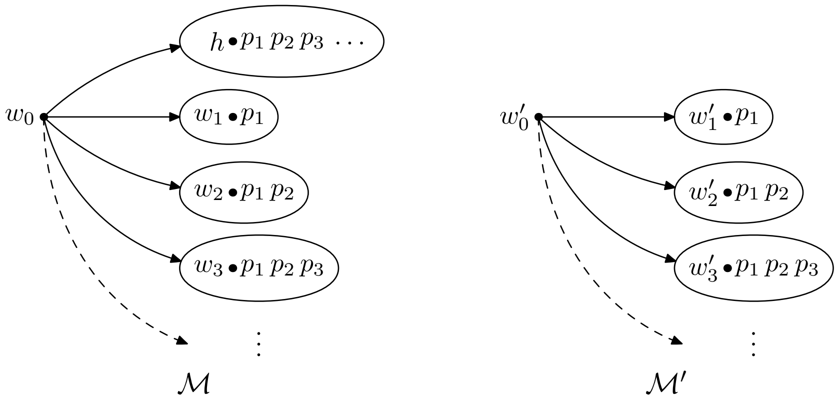

Fig. 3.2: $(\mathcal{M}, w_0)$ y $(\mathcal{M}', w'_0)$ son modalmente equivalentes pero no bisimilares.

Las condiciones de *neighbourhood-finito* e *imagen-finita* son necesarias debido a la finitud de las fórmulas del lenguaje. De igual forma que ocurre en la lógica modal básica con su noción de bisimulación, si agregáramos al lenguaje conjunciones infinitas, las propiedades de *neighbourhood-finito* e *imagen-finita* no serían necesarias como hipótesis de la proposición 3.2, y las nociones de equivalencia modal y bisimilaridad coincidirían. Basta ver que para la construcción del contraejemplo necesario para la demostración de *zig* en la proposición 3.2 podemos usar una conjunción y disyunción posiblemente infinitas de fórmulas de las mismas características que, utilizando el mismo argumento, permiten llegar a un absurdo.

Hemos visto hasta aquí que $\mathcal{N}_{\subseteq}$ tiene una caracterización elegante en términos de una definición de bisimulación adecuada. Veremos en la siguiente sección que la situación es radicalmente diferente para $\mathcal{N}_{=}$.

3.2 Neighbourhood Semantics Strict

En esta sección comienza nuestro aporte original. Definiremos la noción de bisimulación para $\mathcal{N}_{=}$ y para algunas extensiones de interés.

La noción de *Neighbourhood Semantics Strict* fue la introducida originalmente por Montague y Scott como una posible generalización de la semántica relacional de Kripke. Esta es una semántica que tiene también algo de *sintáctico* en su definición, dado que para cada mundo, debo definir exactamente qué *fórmulas* deben ser necesarias. Es decir, si en un mundo w son necesarias las fórmulas $\varphi_1, \ldots, \varphi_n$, la estructura de $\mathcal{N}_{=}$ refleja esto de forma muy transparente: $N_w = \{\|\varphi_1\|, \ldots, \|\varphi_n\|\}$. Esta es exactamente la idea detrás de esta semántica: N mapea mundos en fórmulas que deben ser necesarias.

Pero ¿qué ocurre –teniendo en cuenta lo anterior– si la relación de accesibilidad N_w contiene un elemento X que no es la extensión de ninguna fórmula? Claramente, el elemento X no añadirá ninguna información lógica: no habrán nuevas fórmulas necesarias en w –ni tampoco habrán otras que dejarán de serlo– por la existencia de este elemento. De hecho, la eliminación de este elemento, así como la incorporación de otros elementos como este, no modificará el valor de verdad de ninguna sentencia de la lógica. Estos elementos son, de alguna forma, inocuos. A continuación formalizaremos estas ideas.

3.2.1 Diferenciación

Definición 3.3. Definimos $\mathcal{D}$ como la clase de modelos tales que $\mathcal{M} = \langle W, N, P\rangle \in \mathcal{D}$ sii $\forall w \in W\ \forall T \in N_w\ \exists \varphi_T$ donde $\|\varphi_T\|^{\mathcal{M}} = T$. Llamaremos a $\mathcal{D}$ la clase de modelos *diferenciables* y a φ_T la *fórmula característica* de T.

En otras palabras, $\mathcal{D}$ es la clase de modelos en donde todos los neighbourhoods son necesariamente extensiones de fórmulas.

Definición 3.4. Diremos que T es un conjunto diferenciable en $\mathcal{M}$ sii $\exists \varphi$ tal que $\|\varphi\|^{\mathcal{M}} = T$. Definimos la *operación de diferenciación* sobre un modelo minimal $\mathcal{M} = \langle W, N, P\rangle$, a la que notaremos $\mathcal{M}^d = \langle W, N^d, P\rangle$, que consiste en eliminar todos los conjuntos en N que no tienen fórmula característica. Es decir:

$$N^d_w = \{T \mid T \in N_w \text{ y } T \text{ diferenciable en } \mathcal{M}\}.$$

Como se ve, para cualquier modelo $\mathcal{M}$, resulta que $\mathcal{M}^d \in \mathcal{D}$. Veremos a continuación que esta operación coincide con la intuición introducida antes ya que, efectivamente, estos neighbourhoods eliminados no modifican la relación de satisfacibilidad del modelo.

Proposición 3.3. Sea $\mathcal{M} = \langle W, N, P\rangle$ un modelo minimal. Para cada fórmula modal $\varphi \in \mathcal{N}_=$ y $w \in W$

$$\mathcal{M}, w \models \varphi \iff \mathcal{M}^d, w \models \varphi.$$

Demostración. Por inducción en φ. Como la diferencia entre $\mathcal{M} = \langle W, N, P\rangle$ y $\mathcal{M}^d = \langle W, N^d, P\rangle$ reside en la relación de accesibilidad N, cuando φ es una variable proposicional o una fórmula booleana, la prueba es trivial.

Si $\varphi = \boxminus\psi$:

$\Rightarrow$) Supongamos que $\mathcal{M}, w \models \boxminus\psi$, esto significa que $\|\psi\|^{\mathcal{M}} \in N_w$. Trivialmente, $\|\psi\|^{\mathcal{M}}$ es diferenciable en $\mathcal{M}$. Por la definición 3.4, $\|\psi\|^{\mathcal{M}} \in N^d$ y aplicando la hipótesis inductiva, $\|\psi\|^{\mathcal{M}} = \|\psi\|^{\mathcal{M}^d}$. Entonces, $\mathcal{M}^d, w \models \boxminus\psi$.

$\Leftarrow$) Supongamos que $\mathcal{M}^d, w \models \boxminus\psi$, esto significa que $\|\psi\|^{\mathcal{M}^d} \in N^d_w$. Por la definición 3.4, $\|\psi\|^{\mathcal{M}^d} \in N_w$, y aplicando la hipótesis inductiva, $\|\psi\|^{\mathcal{M}} = \|\psi\|^{\mathcal{M}^d}$. Entonces, $\mathcal{M}, w \models \boxminus\psi$. □

En realidad, preservación vía diferenciación también vale para el lenguaje modal extendido con la modalidad E ($\mathcal{N}_=(\mathsf{E})$). Veremos más adelante que este lenguaje será particularmente interesante para nuestras investigaciones semánticas. Por ahora, establezcamos el resultado de preservación.

Proposición 3.4. Sea $\mathcal{M} = \langle W, N, P\rangle$ un modelo minimal. Para cada fórmula modal $\varphi \in \mathcal{N}_=(\mathsf{E})$ y $w \in W$

$$\mathcal{M}, w \models \varphi \iff \mathcal{M}^d, w \models \varphi.$$

Demostración. Sólo necesitamos considerar el caso adicional para $\varphi = \mathsf{E}\psi$, pero este caso es trivial por hipótesis inductiva. $\mathcal{M}, w \models \mathsf{E}\psi$ sii $\mathcal{M}, w \models \psi$ para algún $v \in W$ sii por hipótesis inductiva $\mathcal{M}^d, v \models \psi$ sii $\mathcal{M}^d, w \models \mathsf{E}\psi$. □

3.2.2 Bisimulación para $\mathcal{N}_=$

A partir de la proposición 3.3 podemos investigar el poder expresivo de $\mathcal{N}_=$ restringiéndonos a modelos diferenciables sin alterar la lógica, que es lo que haremos para definir a continuación la noción de bisimulación. Veremos que las condiciones *zig* y *zag* resultan complejas, y seguidamente analizaremos su significado.

Definición 3.5. Sean $\mathcal{M} = \langle W, N, P\rangle$, $\mathcal{M}' = \langle W', N', P'\rangle$ dos modelos diferenciables. Una *bisimulación diferenciable strict* entre estos modelos es una relación $Z \subseteq W \times W'$ no vacía tal que si xZx' entonces:

1. $\forall p \in \mathsf{PROP}$, $x \in P(p) \iff x' \in P'(p)$
2. (Zig) Para cada T en N_x existe $T' \in N'_{x'}$ tal que:
   a. $\forall w' \in T'\ \exists w \in T$ tal que wZw' y
   b. $\forall X' \notin N'_{x'}$ diferenciable en $\mathcal{M}'$ tal que $T' \subsetneq X' \subseteq W'$, entonces existen elementos $w' \in X' \setminus T', w \in W \setminus T$ que satisfacen wZw'
3. (Zag) Para cada T' en $N'_{x'}$ existe $T \in N_x$ tal que:
   a. $\forall w \in T\ \exists w' \in T'$ tal que wZw' y
   b. $\forall X \notin N_x$ diferenciable en $\mathcal{M}$ tal que $T \subsetneq X \subseteq W$, entonces existen elementos $w \in X \setminus T, w' \in W' \setminus T'$ que satisfacen wZw'

Miremos en mayor detalle lo que dice esta noción de bisimulación. La primera condición está pidiendo armonía atómica, y es claramente necesaria, porque dado que estamos buscando una noción de equivalencia entre modelos que preserve la equivalencia modal, en particular deberá preservar las variables proposicionales. La condición 2 nos pide que todo neighbourhood de x debe estar relacionado con algún neighbourhood de x' de forma tal que se cumplan 2.a y 2.b, y simétricamente la condición 3 exige que cada neighbourhood de x' esté relacionado con un neighbourhood de x de forma tal que se cumplan 3.a y 3.b. Notemos que así como en la noción de bisimulación para la lógica modal básica (definición 1.6) las condiciones *zig* y *zag* exigen que existan ciertas relaciones entre mundos, en este caso –por estar tratando con modelos minimales– se pide la existencia una relación entre *conjuntos de mundos*. De esta forma, las definiciones de bisimulación para *Neighbourhood Semantics* tienen dos niveles de definición: para que dos mundos estén en relación (en bisimulación) debe haber una cierta relación entre sus neighbourhoods, y para que dos neighbourhoods estén en relación, debe cumplirse una serie de relaciones entre los mundos que los componen.

Pero ¿qué dicen exactamente las condiciones 2.a y 2.b (las condiciones 3.a y 3.b son similares)? La condición 2.a exige lo estrictamente necesario para asegurar que si hay un neighbourhood T en donde en *todos* sus puntos se cumple φ, entonces hay un neighbourhood correspondiente T' en el otro modelo en donde también se cumple φ. Para asegurar esto se pide que cualquier punto de T' sea bisimilar a alguno (cualquiera) de T como se muestra gráficamente en la figura 3.3. Esto es exactamente lo mismo que pide la condición *zig* de la definición 3.1 de bisimulación para $\mathcal{N}_\subseteq$.

Sin embargo, 2.a no es suficiente para $\mathcal{N}_=$ porque la semántica asociada a $\boxed{=}$ es más estricta que la asociada a $\boxed{\subseteq}$.

Intuitivamente, podríamos pensar que la condición requerida es el equivalente a 2.a pero sobre el complemento de los neighbourhoods de cada elemento bisimilar. Exploraremos esta

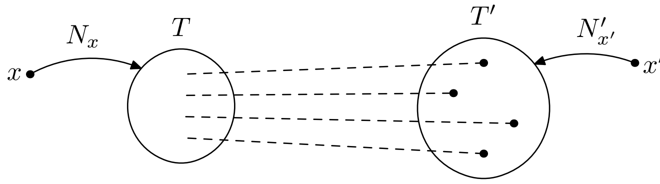

Fig. 3.3: Condición 2.a de bisimulación.

condición y demostraremos en la proposición 3.8, que no es la condición mínima requerida por $\mathcal{N}_=$. Es decir, exhibiremos modelos modalmente equivalentes e intuitivamente bisimilares que no satisfacen la condición de "sincronía sobre complementos". La condición 2.b es más débil y alcanza para asegurar la preservación de equivalencia modal en $\mathcal{N}_=$. Esta condición predica sobre los superconjuntos de T' que no son neighbourhoods de x' (de alguna forma cubre el complemento de $N'_{x'}$), pidiendo que tengan algún elemento fuera de T' bisimilar a algún otro fuera de T, como lo muestra la figura 3.4. El hecho de que X' deba ser diferenciable

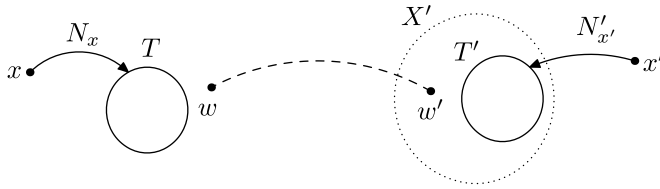

Fig. 3.4: Condición 2.b de bisimulación.

simplemente relaja la condición 2.b, dado que sólo nos interesan los conjuntos de los que podemos hablar, es decir, de aquellos para los que existe una fórmula que los caracterice.

Proposición 3.5. Sean $\mathcal{M} = \langle W, N, P \rangle$, $\mathcal{M}' = \langle W', N', P' \rangle$ dos modelos minimales diferenciables, Z bisimulación minimal strict diferenciable entre ellos, $w \in W$, $w' \in W'$ y wZw'. Entonces, para cada fórmula modal $\varphi \in \mathcal{N}_=$:

$$\mathcal{M}, w \models \varphi \iff \mathcal{M}', w' \models \varphi.$$

Demostración. Por inducción en φ:

<u>Caso base</u>: Si φ es una variable proposicional, la condición 1 de bisimulación asegura que $w \in P(\varphi) \iff w' \in P'(\varphi)$, entonces, $\mathcal{M}, w \models \varphi \iff \mathcal{M}', w' \models \varphi$.

<u>Paso inductivo 1</u>: Si φ es una fórmula booleana de la forma $\neg\psi$ ó $\psi_1 \wedge \psi_2$, sólo tenemos que aplicar la hipótesis inductiva.

<u>Paso inductivo 2</u>: El caso modal: $\varphi = \boxminus\psi$.

$\Rightarrow$) Supongamos que $\mathcal{M}, w \models \boxminus\psi$, entonces $\|\psi\|^{\mathcal{M}} \in N_w$.

Por la condición 2.a de bisimulación debe haber un $X' \in N'_{w'}$ tal que $\forall x' \in X' \; \exists x \in \|\psi\|^{\mathcal{M}}$ tal que xZx'. Sea $x' \in X' \in N'_{w'}$ arbitrario, como cada $x \in \|\psi\|^{\mathcal{M}}$ verifica $\mathcal{M}, x \models \psi$ y además x y x' son bisimilares, podemos aplicar la hipótesis inductiva: $\mathcal{M}', x' \models \psi$. Esto ocurre para todo $x' \in X'$, entonces $X' \subseteq \|\psi\|^{\mathcal{M}'}$.

Para finalmente probar que $\|\psi\|^{\mathcal{M}'} \in N'_{w'}$ primero probaremos que todo mundo en $\|\psi\|^{\mathcal{M}'} \setminus X'$ no es bisimilar a ningún otro de $W \setminus \|\psi\|^{\mathcal{M}}$. Esto es verdadero, porque si alguno de estos mundos s' fuera bisimilar a un mundo $s \in W \setminus \|\psi\|^{\mathcal{M}}$, ocurriría que $\mathcal{M}, s \not\models \psi$ pero $\mathcal{M}', s' \models \psi$, que, por hipótesis inductiva, es absurdo. Si $X' = \|\psi\|^{\mathcal{M}'}$ terminamos, si $X' \subsetneq \|\psi\|^{\mathcal{M}'}$, sabemos por la condición 2.b de bisimulación, que $X' \cup (\|\psi\|^{\mathcal{M}'} \setminus X') \in N'_{w'}$, para $X' \in N'_{w'}$ y todos los elementos de $\|\psi\|^{\mathcal{M}'} \setminus X'$ no son bisimilares a ningún otro de $W \setminus \|\psi\|^{\mathcal{M}}$. Como $X' \subset \|\psi\|^{\mathcal{M}'}$, $X' \cup (\|\psi\|^{\mathcal{M}'} \setminus X') = \|\psi\|^{\mathcal{M}'}$, entonces es diferenciable (por la fórmula ψ), y $\|\psi\|^{\mathcal{M}'} \in N'_{w'}$, o de forma equivalente, $\mathcal{M}', w' \models \boxminus\psi$.

$\Leftarrow$) La vuelta es análoga a $\Rightarrow$) usando las condiciones 3.a y 3.b. □

Como hicimos para $\mathcal{N}_{\subseteq}$, estudiaremos qué condiciones de "finitud" son necesarias para que bisimulación coincida con equivalencia modal. Veremos que en este caso debemos restringirnos a modelos finitos. A diferencia de $\mathcal{N}_{\subseteq}$ en donde necesitamos que el modelo sea neighbourhood-finito e imagen-finita[1] (que no implican modelo finito), en la bisimulación de $\mathcal{N}_{=}$ establecemos condiciones sobre los elementos de cada neighbourhood y sobre los elementos del complemento de cada neighbourhood. Es por eso que necesitamos las propiedades de neighbourhood-finito y co-neighbourhood-finito[2], que implica que el modelo sea finito.

Proposición 3.6. Sean $\mathcal{M} = \langle W, N, P\rangle$, $\mathcal{M}' = \langle W', N', P'\rangle \in \mathcal{D}$ dos modelos finitos diferenciables, $w \in W$, $w' \in W'$ y $\forall\varphi$ fórmula modal $\mathcal{M}, w \models \varphi \iff \mathcal{M}', w' \models \varphi$ (*i.e.*, equivalentes punto a punto). Entonces w y w' son bisimilares (*i.e.*, $\exists Z$ bisimulación minimal strict diferenciable entre $\mathcal{M}$ y $\mathcal{M}'$ tal que wZw').

Demostración. Proponemos que la bisimulación que necesitamos está definida por:

$$xZx' \quad \text{sii} \quad \forall\varphi : \mathcal{M}, x \models \varphi \iff \mathcal{M}', x' \models \varphi$$

Debemos probar que Z es efectivamente una bisimulación. Por definición, $(w, w') \in Z$ lo que asegura que Z es no vacía. Debemos ahora chequear las condiciones 1 a 3 de la definición 3.5.

1. La primera condición de bisimulación es satisfecha por definición de Z.

2. Para probar que la segunda condición es satisfecha, supongamos *ad absurdum* que no lo es, por lo cual w y w' son modalmente equivalentes y en relación wZw', $T \in N_w$ pero $\forall T' \in N'_{w'}$ es falso 2.a o bien es falso 2.b. Es decir, los elementos de $N'_{w'}$ pueden clasificarse en aquellos que falsifican 2.a $\{T'_1, \ldots, T'_t\}$ y aquellos que falsifican 2.b $\{S'_1, \ldots, S'_l\}$. Los conjuntos pueden no ser disjuntos, pero cumplen que $\{T'_1, \ldots, T'_t\} \cup \{S'_1, \ldots, S'_l\} = N'_{w'}$. Además, sean $T = \{h_1, \ldots, h_n\}$ y $W \setminus T = \{\bar{h}_1, \ldots, \bar{h}_m\}$.

   Sea $\psi_0 := \varphi_T$ donde φ_T es la fórmula característica de T. Como $\mathcal{M}, w \models \boxminus\psi_0$, entonces $\mathcal{M}', w' \models \boxminus\psi_0$, así que $\|\psi_0\|^{\mathcal{M}'} \in N'_{w'}$. Luego repitamos el siguiente argumento de razonamiento teniendo como invariante que $\|\psi_r\|^{\mathcal{M}'} \in N'_{w'}$. Veremos que en cada iteración, $\|\psi_r\|^{\mathcal{M}'} \subsetneq \|\psi_{r+1}\|^{\mathcal{M}'}$. En una cantidad finita de pasos llegaremos a un ψ_q tal que $\mathcal{M}, w \models \boxminus\psi_q$ pero $\mathcal{M}', w' \not\models \boxminus\psi_q$, lo cual es absurdo.

   En el paso r:

[1] Cf. proposición 3.2 para $\mathcal{N}_{\subseteq}$.
[2] Ver definición 3.2.

- Si $\|\psi_{r-1}\|^{\mathcal{M}'} = T'_k$. Entonces existe $h'_k \in T'_k$ tal que para cada $h_i \in T$ hay un ψ_i^k donde $\mathcal{M}, h_i \models \psi_i^k$ pero $\mathcal{M}', h'_k \not\models \psi_i^k$. Sea $\psi_r = \psi_{r-1} \wedge (\bigvee_{j=1}^{n} \psi_j^k)$. Se ve fácilmente que $\mathcal{M}, w \models \boxminus\psi_r$. Si $\mathcal{M}', w' \not\models \boxminus\psi_r$, absurdo. Si no, iteremos teniendo en cuenta que $\|\psi_r\|^{\mathcal{M}'} \subsetneq \|\psi_{r-1}\|^{\mathcal{M}'}$ (porque h'_k ya no satisface ψ_r) y por lo tanto $\#\|\psi_r\|^{\mathcal{M}'} < \#\|\psi_{r-1}\|^{\mathcal{M}'}$, y además $\|\psi_{r-1}\|^{\mathcal{M}} = \|\psi_r\|^{\mathcal{M}} = T$.
- Si $\|\psi_{r-1}\|^{\mathcal{M}'} = S'_k$. Entonces, existe $\bar{X}' = S'_k \cup \{\bar{x}'_1, \ldots, \bar{x}'_z\} \notin N'_{w'}$ tal que $S'_k \cap \{\bar{x}'_1, \ldots, \bar{x}'_z\} = \emptyset$, diferenciable por su fórmula característica (llamémosla η), donde para cada $\bar{x}'_i \in \bar{X}' \setminus S'_k$ y $\bar{h}_j \in W \setminus T$ hay un $\bar{\psi}_{i,j}^k$ en donde $\mathcal{M}, \bar{h}_j \models \bar{\psi}_{i,j}^k$ pero $\mathcal{M}', \bar{x}'_i \not\models \bar{\psi}_{i,j}^k$.
  Sea $\psi_r = \psi_{r-1} \vee (\eta \wedge \bigvee_{i=1}^{z} \bigwedge_{j=1}^{m} \neg\bar{\psi}_{i,j}^k)$.
  Como no hay un elemento $x \in W \setminus T$ donde $\mathcal{M}, x \models \bigwedge_{j=0}^{m} \neg\bar{\psi}_{i,j}^k$ para ningún i, entonces no hay ningún elemento $x \in W \setminus T$ donde $\mathcal{M}, x \models \bigvee_{i=1}^{z} \bigwedge_{j=1}^{m} \neg\bar{\psi}_{i,j}^k$. Esto significa que $\|\psi_{r-1}\|^{\mathcal{M}} = \|\psi_r\|^{\mathcal{M}} = T$.
  Pero cada $x' \in \|\psi_{r-1}\|^{\mathcal{M}'}$ verifica $\mathcal{M}', x' \models \psi_r$, y $\|\eta \wedge \bigvee_{i=1}^{z} \bigwedge_{j=1}^{m} \neg\bar{\psi}_{i,j}^k\|^{\mathcal{M}'} = X'$. Entonces $\|\psi_r\|^{\mathcal{M}'} = X'$. Y como mencionamos antes, $\|\psi_r\|^{\mathcal{M}'} \notin N'_{w'}$.
  En este momento podemos parar de iterar dado que tenemos que $\mathcal{M}, w \models \boxminus\psi_r$ pero $\mathcal{M}', w' \not\models \boxminus\psi_r$. Absurdo.

3. La tercera condición se prueba por analogía con la segunda condición.

□

Lema 3.3. La proposición 3.6 no es válida para modelos minimales sin la propiedad de neighbourhood-finito.

Demostración. Utilizando el mismo modelo minimal que en el lema 3.1 (figura 3.1), es fácil ver que h' no es modalmente equivalente a *ninguno* de los elementos en N_{w_0}, por la misma razón que en el lema 3.1, dado que, como ya hemos mencionado, la condición 3 de la definición 3.1 para $\mathcal{N}_{\subseteq}$ es la misma que la condición 3.a de la definición 3.5 para $\mathcal{N}_{=}$. Al no ser satisfecha la condición 3.a, w_0 y w'_0 no son bisimilares y sin embargo, son modalmente equivalentes. □

Lema 3.4. La proposición 3.6 no es válida para modelos minimales sin la propiedad de ser co-neighbourhood-finito[3].

Demostración. El contraejemplo es similar al usado en el lema 3.3 pero todos los elementos deben ser posicionados *afuera* del neighbourhood como lo muestra la figura 3.5. Alcanza entonces con tener $N_{w_0} = N'_{w'_0} = \emptyset$. Aquí, la condición que no es satisfecha es 3.b.

□

Corolario 3.1. La proposición 3.6 no es válida para modelos minimales infinitos.

Demostración. Trivial de los lemas 3.3 y 3.4. □

En este punto, debemos notar que con la inclusión de fórmulas con conjunciones infinitas, la proposición 3.6 no es trivialmente válida para modelos posiblemente infinitos. Esto se debe a que en este caso la demostración utiliza un argumento iterativo que es sólo válido para conjuntos finitos.

[3] Ver definición 3.2

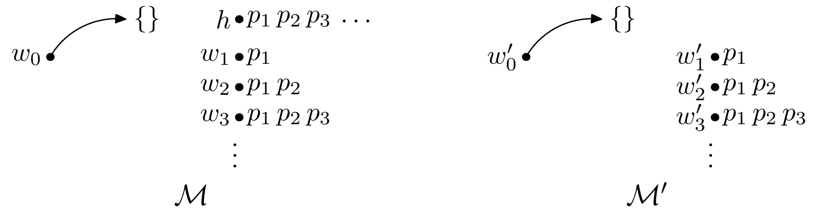

Fig. 3.5: $(\mathcal{M}, w_0)$ y $(\mathcal{M}', w'_0)$ son modalmente equivalentes pero no bisimilares.

Definición 3.6. Sean $\mathcal{M} = \langle W, N, P\rangle$, $\mathcal{M}' = \langle W', N', P'\rangle$ dos modelos minimales. Definimos una *bisimulación minimal strict* entre $\mathcal{M}$ y $\mathcal{M}'$ como una bisimulación minimal strict diferenciable entre $\mathcal{M}^d$ y $\mathcal{M}'^d$.

Proposición 3.7. Las proposiciones 3.5 y 3.6 son válidas en la clase de todos los modelos minimales, bajo bisimulación minimal strict.

Demostración. Resulta inmediato de la proposición 3.3 (p. 25) y la definición 3.6. □

Hasta ahora, obtuvimos la noción que parece ser la adecuada para $\mathcal{N}_=$, pero resulta extraño que sea tan compleja, mientras que en el caso de $\mathcal{N}_\subseteq$ la noción resulta natural. A lo largo del estudio de la noción de bisimulación, observamos que el punto clave de su complejidad radica en tener o no la habilidad de distinguir superconjuntos. En la siguiente sección investigaremos este tema.

3.2.3 El operador $\triangledown$

Recordemos que durante la discusión acerca del significado de las condiciones *zig* y *zag* de la definición 3.5 de bisimulación hemos considerado una definición alternativa de bisimulación que resulta más natural. Retomemos esta noción de bisimulación intuitiva y comprobemos mediante un contraejemplo que no es la definición adecuada para $\mathcal{N}_=$.

Definición 3.7. Sean $\mathcal{M} = \langle W, N, P\rangle$, $\mathcal{M}' = \langle W', N', P'\rangle$ dos modelos minimales. Una *bisimulación minimal diferenciable simétrica strict* entre estos modelos es una relación $Z \subseteq W \times W'$ no vacía tal que si xZx' entonces:

1. $\forall p$ variable proposicional, $x \in P(p) \iff x' \in P'(p)$
2. (Zig) Para cada $T \in N_x$ debe haber un conjunto $T' \in N'_{x'}$ tal que:

   a. $\forall w' \in T'\ \exists w \in T$ tal que wZw' y b. $\forall w' \in W' \setminus T'\ \ \exists w \in W \setminus T$ tal que wZw'

3. (Zag) Para cada $T' \in N'_{x'}$ debe haber un conjunto $T \in N_x$ tal que:

   a. $\forall w \in T\ \exists w' \in T'$ tal que wZw' y b. $\forall w \in W \setminus T\ \ \exists w' \in W' \setminus T'$ tal que wZw'

La definición 3.7, que tiene la condición simétrica de la definición 3.1 y parece mucho más natural, resulta demasiado fuerte para $\mathcal{N}_=$.

Proposición 3.8. Existen modelos minimales finitos y diferenciables, modalmente equivalentes pero no bisimilares bajo la definición 3.7.

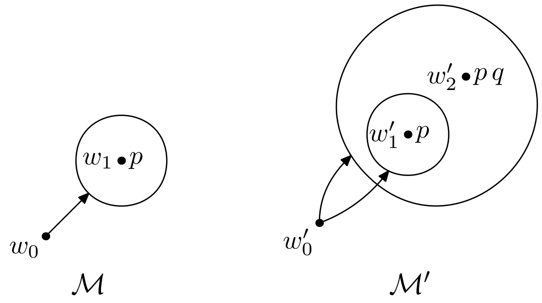

Fig. 3.6: $(\mathcal{M}, w_0)$ y $(\mathcal{M}', w'_0)$ son modalmente equivalentes pero no bisimilares.

Demostración. Veamos un contraejemplo ilustrador:

Sea $\mathcal{M} = \langle W, N, P\rangle$, donde $W = \{w_0, w_1\}$, $P(p) = \{w_1\}$, $N_{w_0} = \{\{w_1\}\}$.

Y sea $\mathcal{M}' = \langle W', N', P'\rangle$, donde $W' = \{w'_0, w'_1, w'_2\}$, $P'(p) = \{w'_1, w'_2\}$, $P'(q) = \{w'_2\}$, $N'_{w'_0} = \{\{w'_1\}, \{w'_1, w'_2\}\}$ como se muestra en la figura 3.6.

Notemos que estos modelos son diferenciables, el único neighbourhood en $\mathcal{M}$ puede ser expresado por $\|p\|^{\mathcal{M}}$ y los neighbourhoods de $\mathcal{M}'$ por $\|p\|^{\mathcal{M}'}$ y $\|p \wedge \neg q\|^{\mathcal{M}'}$. Estos dos modelos son modalmente equivalentes pero no son bisimilares, y sin embargo hubiéramos esperado (siendo finitos y diferenciables) que fueran bisimilares. $\square$

Esta bisimulación simétrica falla cuando hay dos neighbourhoods incluidos uno en el otro. La lógica $\mathcal{N}_=$ no puede ver la diferencia entre un modelo con un sólo neighbourhood y otro con dos neighbourhoods, donde uno es un subconjunto propio del otro. Es decir, mientras que la equivalencia modal colapsa superconjuntos, la bisimulación simétrica de la definición 3.7 percibe estas diferencias. Podríamos entonces extender el lenguaje, agregando un operador que ayudara a *ver* (distinguir) estos superconjuntos. La lógica extendida podrá, por ejemplo, diferenciar entre $\mathcal{M}$ y $\mathcal{M}'$ de la figura 3.6.

Vamos a enriquecer el lenguaje para poder entender qué expresividad le falta a $\mathcal{N}_=$ para tener la bisimulación simétrica recién vista. Vale recalcar que nuestro interés es caracterizar específicamente esta expresividad ausente en $\mathcal{N}_=$, sin agregar nada "de más", es decir, queremos agregar la *mínima* expresividad necesaria para llegar exactamente a la bisimulación simétrica de la definición 3.7. De esta forma tendremos una herramienta para entender mejor las diferencias "semánticas" entre las dos bisimulaciones[4] que hemos considerado para $\mathcal{N}_=$.

Definición 3.8. Introducimos un nuevo operador binario que denotaremos "$\triangledown$", con la siguiente semántica:

$$\mathcal{M}, w \models \varphi \triangledown \psi \quad \text{sii} \quad \exists X_1, X_2 \in N_w \text{ tal que } X_1 \neq X_2, X_1 = \|\varphi\|^{\mathcal{M}}, X_2 = \|\varphi \vee \psi\|^{\mathcal{M}}.$$

Por consiguiente definiremos el lenguaje $\mathcal{N}_=(\triangledown)$ inductivamente como:

$$\mathcal{N}_=(\triangledown) := p \mid \neg\varphi \mid \varphi \wedge \varphi' \mid \boxminus\varphi \mid \varphi \triangledown \varphi'$$

donde $p \in \mathsf{PROP}$ y $\varphi, \varphi' \in \mathcal{N}_=(\triangledown)$.

[4] Cf. definición 3.7 con la definición 3.5 de bisimulación para $\mathcal{N}_=$.

Proposición 3.9. Sean $\mathcal{M} = \langle W, N, P\rangle$, $\mathcal{M}' = \langle W', N', P'\rangle$ dos modelos minimales, Z una bisimulación minimal strict simétrica diferenciable entre ellos, $w \in W$, $w' \in W'$ y wZw'. Entonces, para cada fórmula modal $\varphi \in \mathcal{N}_{=}(\triangledown)$:

$$\mathcal{M}, w \models \varphi \iff \mathcal{M}', w' \models \varphi$$

Demostración. Para demostrar esta proposición, necesitaremos hacer inducción sobre la fórmula, siendo $\triangledown > \boxminus$. Con ese fin redefiniremos la función de complejidad de las fórmulas $|\cdot| : \mathcal{N}_{=}(\triangledown) \to \mathbb{N}$ de la siguiente manera:

$$\begin{array}{rcl} |p| & = & 0 \\ |\neg\varphi| & = & 1 + |\varphi| \\ |\varphi \wedge \psi| & = & 1 + |\varphi| + |\psi| \\ |\boxminus\varphi| & = & 1 + |\varphi| \\ |\varphi\triangledown\psi| & = & 3 + |\varphi| + |\psi| \end{array}$$

Por inducción en la complejidad de φ:

Caso base: Si φ es una variable proposicional, la condición 1 de bisimulación asegura que $w \in P(\varphi) \iff w' \in P'(\varphi)$, entonces $\mathcal{M}, w \models \varphi \iff \mathcal{M}', w' \models \varphi$.

Paso inductivo 1: Si φ es una fórmula booleana de la forma $\neg\psi$ ó $\psi_1 \wedge \psi_2$, sólo resta aplicar la hipótesis inductiva.

Paso inductivo 2: La modalidad unaria: $\varphi = \boxminus\psi$

$\Rightarrow$) Supongamos que $\mathcal{M}, w \models \boxminus\psi$, entonces $\|\psi\|^{\mathcal{M}} \in N_w$.

Por la condición 2.a de bisimulación debe haber un $X' \in N'_{w'}$ de forma tal que $\forall x' \in X'\ \exists x \in X$ tal que xZx'. Sea $x' \in X' \in N'_{w'}$ arbitrario, como cada $x \in X$ verifica $\mathcal{M}, x \models \psi$ y además x y x' son bisimilares, podemos aplicar la hipótesis inductiva: $\mathcal{M}', x' \models \psi$. Esto ocurre para todo $x' \in X'$, entonces $X' \subseteq \|\psi\|^{\mathcal{M}'}$.

Por la condición 2.b de bisimulación, cada $t' \in W' \setminus X'$ debe ser bisimilar a un elemento de $W \setminus \|\psi\|^{\mathcal{M}}$, que es el conjunto de todos los elementos t tales que $\mathcal{M}, t \not\models \psi$. Aplicando la hipótesis inductiva, $\mathcal{M}', t' \not\models \psi$. Como esto ocurre para cada $t' \in W' \setminus X'$, entonces $\|\psi\|^{\mathcal{M}'} \subseteq X'$.

Podemos entonces concluir que $\|\psi\|^{\mathcal{M}'} = X'$, o equivalentemente, $\mathcal{M}', w' \models \boxminus\psi$.

$\Leftarrow$) Análogo a $\Rightarrow$).

Paso inductivo 3: La modalidad binaria: $\varphi = \psi_1\triangledown\psi_2$

$\Rightarrow$) Supongamos que $\mathcal{M}, w \models \psi_1\triangledown\psi_2$, entonces $\exists X_1, X_2 \in N_w$ tal que $X_1 \subsetneq X_2$, $X_1 = \|\psi_1\|^{\mathcal{M}}$, $X_2 = \|\psi_1 \vee \psi_2\|^{\mathcal{M}}$. Como esto implica que $\mathcal{M}, w \models \boxminus\psi_1$ y $\mathcal{M}, w \models \boxminus(\psi_1 \vee \psi_2)$, podemos aplicar la hipótesis inductiva dado que

$$|\boxminus(\psi_1 \vee \psi_2)| = |\psi_1 \vee \psi_2| + 1 = |\psi_1| + |\psi_2| + 2 < |\psi_1| + |\psi_2| + 3 = |\psi_1\triangledown\psi_2|$$

y entonces saber que existe $X'_1, X'_2 \in N'_{w'}$ tal que $X'_1 = \|\psi_1\|^{\mathcal{M}'}$ y $X'_2 = \|\psi_1 \vee \psi_2\|^{\mathcal{M}'}$.

Todavía nos falta probar que $X'_1 \neq X'_2$, en otras palabras, que existe $t' \in X'_2 \setminus X'_1$. Sea $t \in X_2 \setminus X_1$. Debido a la condición 3.b debe existir $X''_1 \in N_w$ donde todos los elementos de $W \setminus X''_1$ son bisimilares a alguno de $W' \setminus X'_1$. En particular, t. Notemos que $X''_1 \subseteq X_1$, de otra forma habría un elemento $z \notin X_1 = \|\psi_1\|^{\mathcal{M}}$, $z \in X_1''$ que debe ser bisimilar (por la condición 3.a) a un elemento $z' \in X_1' = \|\psi_1\|^{\mathcal{M}'}$, lo cual, por hipótesis inductiva, es absurdo.

Ahora bien, podría ocurrir que la imagen bisimilar t' esté en $X_2' \setminus X_1'$ o que esté en $W' \setminus X_2'$. Es fácil ver que este último caso no es posible, porque significaría que $\mathcal{M}', t' \not\models \psi_2$, y aplicando la hipótesis inductiva esto implica que $\mathcal{M}, t \not\models \psi_2$, que es absurdo dado que $t \in X_2 \setminus X_1 = \|\psi_2\|^{\mathcal{M}}$. De esta forma, t' está en $X_2' \setminus X_1'$ como requeríamos.

$\Leftarrow$) Análogo a $\Rightarrow$). $\square$

Comprobemos una vez más en qué casos equivalencia modal implica bisimulación.

Proposición 3.10. Sean $\mathcal{M} = \langle W, N, P\rangle$, $\mathcal{M}' = \langle W', N', P'\rangle \in \mathcal{D}$ dos modelos minimales finitos diferenciables, $w \in W$, $w' \in W'$, y $\forall\varphi \in \mathcal{N}_{=}(\nabla)$ $\mathcal{M}, w \models \varphi \iff \mathcal{M}', w' \models \varphi$ (*i.e.*, equivalentes punto a punto). Entonces w y w' son bisimilares (*i.e.*, $\exists Z$ bisimulación minimal strict simétrica diferenciable entre $\mathcal{M}$ y $\mathcal{M}'$ tal que wZw').

Demostración. La bisimulación que necesitamos está definida por:

xZx' sii $\forall\varphi : \mathcal{M}, x \models \varphi \iff \mathcal{M}', x' \models \varphi$

Necesitamos probar que esta es efectivamente una bisimulación. Por definición, $(w, w') \in Z$ lo que asegura que Z es no vacía. Debemos ahora chequear las condiciones 1 a 3 de la definición 3.7.

1. La primera condición de bisimulación es satisfecha por definición.

2. Para chequear que la segunda condición es satisfecha, supongamos *ad absurdum* que no lo es, es decir que w y w' son modalmente equivalente y en relación wZw', $T \in N_w$ pero $\forall T' \in N'_{w'}$ es falso 2.a o es falso 2.b. Sea $N'_{w'} = \{T_1', \ldots, T_t'\} \cup \{S_1', \ldots, S_l'\}$ donde el primer conjunto falsifica la condición 2.a y el segundo falsifica 2.b, $T = \{h_1, \ldots, h_n\}$ y $W \setminus T = \{\bar{h}_1, \ldots, \bar{h}_m\}$.

   Inicializar $\psi_0 = \varphi_T$ donde φ_T es la fórmula característica de T. Como $\mathcal{M}, w \models \boxminus\psi_0$, entonces $\mathcal{M}', w' \models \boxminus\psi_0$, así que $\|\psi_0\|^{\mathcal{M}'} \in N'_{w'}$. Luego repetir el siguiente argumento de razonamiento teniendo como invariante que $\|\psi_r\|^{\mathcal{M}'} \in N'_{w'}$. Veremos que en cada iteración, $\|\psi_r\|^{\mathcal{M}'} \subsetneq \|\psi_{r+1}\|^{\mathcal{M}'}$. En una cantidad finita de pasos llegaremos a un ψ_q tal que $\mathcal{M}, w \models \boxminus\psi_q$ pero $\mathcal{M}', w' \not\models \boxminus\psi_q$, lo cual es absurdo.

   En el paso r:

   - Si $\|\psi_{r-1}\|^{\mathcal{M}'} = T_k'$. Entonces debe existir $h_k' \in T_k'$ tal que para cada $h_i \in T$ hay un ψ_i^k donde $\mathcal{M}, h_i \models \psi_i^k$ pero $\mathcal{M}', h_k' \models \psi_i^k$. Sea $\psi_r = \psi_{r-1} \wedge (\bigvee_{j=1}^{n} \psi_j^k)$. Diremos que $\mathcal{M}, w \models \boxminus\psi_r$. Si $\mathcal{M}', w' \not\models \boxminus\psi_r$, absurdo. Si no, iterar notando que $\|\psi_r\|^{\mathcal{M}'} \subsetneq \|\psi_{r-1}\|^{\mathcal{M}'}$ (porque h_k' ya no satisface ψ_r) y por lo tanto $\#\|\psi_r\|^{\mathcal{M}'} < \#\|\psi_{r-1}\|^{\mathcal{M}'}$, y además $\|\psi_{r-1}\|^{\mathcal{M}} = \|\psi_r\|^{\mathcal{M}} = T$.
   - Si $\|\psi_{r-1}\|^{\mathcal{M}'} = S_k'$. Entonces debe existir $\bar{h}' \in W' \setminus S_k'$ tal que para cada $\bar{h}_i \in W \setminus T$ hay un $\bar{\psi}_i^k$ donde $\mathcal{M}, \bar{h}_i \not\models \bar{\psi}_i^k$ pero $\mathcal{M}', \bar{h}' \models \bar{\psi}_i^k$.
     Sea $\varphi^k = \bigwedge_{i=0}^{m} \bar{\psi}_i^k$, $\psi_r = \psi_{r-1} \vee \varphi^k$:
     - Si $\|\psi_r\|^{\mathcal{M}'} \notin N'_{w'}$, entonces $\mathcal{M}, w \models \boxminus\psi_r$ pero $\mathcal{M}', w' \not\models \boxminus\psi_r$. Absurdo.
     - En otro caso, $\mathcal{M}', w' \models \psi_{r-1} \nabla \varphi^k$ pero $\mathcal{M}, w \not\models \psi_{r-1} \nabla \varphi^k$. Absurdo.

3. La tercera condición es probada por analogía con la segunda condición.

□

Utilizando los mismos contraejemplos que los empleados para $\mathcal{N}_=$ en los lemas 3.3 y 3.4, se demuestra que la condición de modelo *finito* como hipótesis de la proposición 3.10 es efectivamente necesaria.

Definición 3.9. Sean $\mathcal{M} = \langle W, N, P\rangle$, $\mathcal{M}' = \langle W', N', P'\rangle$ dos modelos minimales. Definimos una *bisimulación minimal strict simétrica* entre $\mathcal{M}$ y $\mathcal{M}'$ como una bisimulación minimal strict simétrica diferenciable entre $\mathcal{M}^d$ y $\mathcal{M}'^d$.

Corolario 3.2. Sea $\mathcal{M} = \langle W, N, P\rangle$ un modelo minimal. Para cada fórmula modal $\varphi \in \mathcal{N}_=(\triangledown)$ y $w \in W$

$$\mathcal{M}, w \models \varphi \iff \mathcal{M}^d, w \models \varphi$$

Demostración. Por inducción en la complejidad de la fórmula, definida de la misma forma que en la demostración de la proposición 3.9. Los casos booleanos y el caso de $\Box\psi$ son iguales que en la proposición 3.3.

Consideremos el caso de $\varphi = \psi_1 \triangledown \psi_2$. Esto es trivial de la hipótesis inductiva, $\mathcal{M}, w \models \psi_1 \triangledown \psi_2$ sii existen X_1 y X_2 en N_w tal que $X_1 \neq X_2$ y $\|\psi_1\|^{\mathcal{M}} = X_1$, $\|\psi_1 \vee \psi_2\|^{\mathcal{M}} = X_2$ sii, por hipótesis inductiva, $\|\psi_1\|^{\mathcal{M}^d} = X_1$, $\|\psi_1 \vee \psi_2\|^{\mathcal{M}^d} = X_2$ sii, como ambos neighbourhoods con claramente diferenciables, $\mathcal{M}^d, w \models \psi_1 \triangledown \psi_2$. □

Proposición 3.11. Las proposiciones 3.9 y 3.10 son válidas en la clase de todos los modelos minimales bajo bisimulaciones minimales strict simétricas.

Demostración. Resulta inmediato desde el corolario 3.2 y la definición 3.9. □

En esta sección hemos comprobado que usando $\triangledown$ obtenemos una bisimulación natural. Este resultado nos abre camino para considerar extensiones conocidas que contengan la expresividad de este operador. Estas extensiones tienen especial interés, porque sus bisimulaciones partirán como base de la bisimulación natural e intuitiva de la definición 3.7. El operador $\triangledown$, aunque posiblemente interesante por sí mismo, podría considerarse un operador "artificial" diseñado especialmente para obtener la proposición 3.10. Entonces, ¿podemos definir $\triangledown$ en un lenguaje clásico más expresivo que $\mathcal{N}_=$? En la siguiente sección veremos que la respuesta es *sí*, utilizando la extensión de $\mathcal{N}_=$ con el operador universal.

3.2.4 Generalización para $\mathcal{N}_=(\mathsf{E})$

Definición 3.10. La modalidad existencial E es definida de la siguiente forma:

$$\mathcal{M}, w \models \mathsf{E}\varphi \quad \text{sii} \quad \exists v \in W \text{ tal que } \mathcal{M}, v \models \varphi.$$

Definiremos entonces el lenguaje

$$\mathcal{N}_=(\mathsf{E}) := p \mid \neg\varphi \mid \varphi \wedge \varphi' \mid \boxminus\varphi \mid \mathsf{E}\varphi$$

donde $p \in \mathsf{PROP}$ y $\varphi, \varphi' \in \mathcal{N}_=(\mathsf{E})$.

Ahora bien, consideremos la definición 3.8 donde damos la semántica de "$\triangledown$". Podemos ver que $\triangledown$ es expresable usando sólo $\boxminus$ y la modalidad existencial E que nos da acceso a predicar sobre todo el dominio del modelo.

$$\mathcal{M}, w \models \varphi \triangledown \psi \quad \text{sii} \quad \mathcal{M}, w \models \boxminus\varphi \wedge \boxminus(\varphi \vee \psi) \wedge \mathsf{E}(\psi \wedge \neg\varphi).$$

Es lógico entonces considerar que la bisimulación para $\mathcal{N}_{=}(\mathsf{E})$ puede ser definida como se muestra a continuación.

Definición 3.11. Sean $\mathcal{M} = \langle W, N, P\rangle$, $\mathcal{M}' = \langle W', N', P'\rangle$ dos modelos minimales. Una *bisimulación minimal strict total simétrica diferenciable* entre estos modelos es una relación $Z \subseteq W \times W'$ no vacía tal que $W = \mathrm{Dom}(Z)$, $W' = \mathrm{Im}(Z)$, y siempre que xZx', entonces:

1. $\forall p$ variable proposicional, $x \in P(p) \iff x' \in P'(p)$
2. (Zig) Para cada $T \in N_x$ debe haber otro conjunto $T' \in N'_{x'}$ tal que:

   a. $\forall w' \in T'\ \exists w \in T$ tal que wZw' y b. $\forall w' \in W' \setminus T'\ \ \exists w \in W \setminus T$ tal que wZw'

3. (Zag) Para cada $T' \in N'_{x'}$ debe haber otro conjunto $T \in N_x$ tal que:

   a. $\forall w \in T\ \exists w' \in T'$ tal que wZw' y b. $\forall w \in W \setminus T\ \ \exists w' \in W' \setminus T'$ tal que wZw'

Proposición 3.12. Sean $\mathcal{M} = \langle W, N, P\rangle$, $\mathcal{M}' = \langle W', N', P'\rangle$ dos modelos minimales, Z una bisimulación minimal strict total simétrica diferenciable entre ellos, $w \in W$, $w' \in W'$ y wZw'. Entonces, para cada fórmula modal $\varphi \in \mathcal{N}_{=}(\mathsf{E})$:

$$\mathcal{M}, w \models \varphi \iff \mathcal{M}', w' \models \varphi$$

Demostración. Por inducción en φ:

<u>Caso base</u>: Si φ es una variable proposicional, la condición 1 de bisimulación asegura que $w \in P(\varphi) \iff w' \in P'(\varphi)$, entonces $\mathcal{M}, w \models \varphi \iff \mathcal{M}', w' \models \varphi$.

<u>Paso inductivo 1</u>: Si φ es una fórmula booleana de la forma $\neg\psi$ ó $\psi_1 \wedge \psi_2$, sólo resta aplicar la hipótesis inductiva.

<u>Paso inductivo 2</u>: El caso modal básico: $\varphi = \boxminus\psi$.

$\Rightarrow$) Supongamos que $\mathcal{M}, w \models \boxminus\psi$, entonces $\|\psi\|^{\mathcal{M}} \in N_w$.
Por la condición 2.a de bisimulación debe haber $X' \in N'_{w'}$ de tal forma que $\forall x' \in X'\ \exists x \in X$ tal que xZx'. Sea $x' \in X' \in N'_{w'}$ arbitrario, como cada $x \in X$ verifica $\mathcal{M}, x \models \psi$ y además x y x' son bisimilares, podemos aplicar la hipótesis inductiva: $\mathcal{M}', x' \models \psi$. Esto ocurre para todo $x' \in X'$, entonces $X' \subseteq \|\psi\|^{\mathcal{M}'}$.
Por la condición 2.b de bisimulación cada $t' \in W' \setminus X'$ debe ser bisimilar a un elemento de $W \setminus \|\psi\|^{\mathcal{M}}$, que es el conjunto de todos los elementos t tales que $\mathcal{M}, t \not\models \psi$. Aplicando la hipótesis inductiva, $\mathcal{M}', t' \not\models \psi$. Como esto ocurre para todo $t' \in W' \setminus X'$, entonces $\|\psi\|^{\mathcal{M}'} \subseteq X'$.
Podemos entonces concluir que $\|\psi\|^{\mathcal{M}'} = X'$, o equivalentemente, $\mathcal{M}', w' \models \boxminus\psi$.
$\Leftarrow$) Análogo a $\Rightarrow$).

<u>Paso inductivo 3</u>: El caso modal universal: $\varphi = \mathsf{E}\psi$.

$\Rightarrow$) Supongamos que $\mathcal{M}, w \models \mathsf{E}\psi$, entonces $\exists x \in W$ donde $\mathcal{M}, x \models \psi$. Como la bisimulación entre $\mathcal{M}$ y $\mathcal{M}'$ debe ser *total*, existe $x' \in W'$ bisimilar a x. Podemos entonces aplicar la hipótesis inductiva y afirmar que $\mathcal{M}', x' \models \psi$. Entonces, $\mathcal{M}', w' \models E\psi$.

$\Leftarrow$) Análogo a $\Rightarrow$). $\square$

Proposición 3.13. Sean $\mathcal{M} = \langle W, N, P\rangle$, $\mathcal{M}' = \langle W', N', P'\rangle \in \mathcal{D}$ dos modelos minimales diferenciables finitos, $w \in W$, $w' \in W'$ y $\forall\varphi \in \mathcal{N}_{=}(\mathsf{E}) : \mathcal{M}, w \models \varphi \iff \mathcal{M}', w' \models \varphi$ (*i.e.*, son equivalentes punto a punto). Entonces w y w' son bisimilares (*i.e.*, $\exists Z$ bisimulación minimal strict total simétrica diferenciable entre $\mathcal{M}$ y $\mathcal{M}'$ tal que wZw').

Demostración. Afirmamos que la bisimulación que necesitamos está definida por:

$$xZx' \text{ sii } \forall\varphi : \mathcal{M}, x \models \varphi \iff \mathcal{M}', x' \models \varphi.$$

Debemos probar que esta es efectivamente una bisimulación. Por definición, $(w, w') \in Z$ lo que asegura que Z es no vacía. Debemos ahora chequear las condiciones 1 a 3 de la definición 3.11.

1. La primera condición de bisimulación es satisfecha por definición.

2. Para probar que la segunda condición es satisfecha, supongamos *ad absurdum* que no lo es, implicando esto que w y w' son modalmente equivalentes y en relación wZw', $T \in N_w$ pero $\forall T' \in N'_{w'}$ es falso 2.a o es falso 2.b. Sea $N'_{w'} = \{T'_1, \ldots, T'_t\} \cup \{S'_1, \ldots, S'_l\}$ donde los elementos del primer conjunto falsifican la condición 2.a y los del segundo conjunto falsifican la condición 2.b, $T = \{h_1, \ldots, h_n\}$ y $W \setminus T = \{\bar{h}_1, \ldots, \bar{h}_m\}$.

   Inicialicemos $\psi_0 = \varphi_T$ donde φ_T es la fórmula característica de T. Como $\mathcal{M}, w \models \boxminus\psi_0$, entonces $\mathcal{M}', w' \models \boxminus\psi_0$, así que $\|\psi_0\|^{\mathcal{M}'} \in N'_{w'}$. Luego repitamos el siguiente argumento de razonamiento teniendo como invariante que $\|\psi_r\|^{\mathcal{M}'} \in N'_{w'}$. Veremos que en cada iteración, $\|\psi_r\|^{\mathcal{M}'} \subsetneq \|\psi_{r+1}\|^{\mathcal{M}'}$. En una cantidad finita de pasos llegaremos a un ψ_q tal que $\mathcal{M}, w \models \boxminus\psi_q$ pero $\mathcal{M}', w' \not\models \boxminus\psi_q$, lo cual es absurdo.

   En el paso r:

   - Si $\|\psi_{r-1}\|^{\mathcal{M}'} = T'_k$. Luego existe $h'_k \in T'_k$ tal que para cada $h_i \in T$ hay un ψ_i^k donde $\mathcal{M}, h_i \models \psi_i^k$ pero $\mathcal{M}', h'_k \not\models \psi_i^k$. Sea $\psi_r = \psi_{r-1} \wedge (\bigvee_{j=1}^{n} \psi_j^k)$. Afirmamos que $\mathcal{M}, w \models \boxminus\psi_r$. Si $\mathcal{M}', w' \not\models \boxminus\psi_r$, absurdo. Si no, iteremos notando que $\|\psi_r\|^{\mathcal{M}'} \subsetneq \|\psi_{r-1}\|^{\mathcal{M}'}$ (porque h'_k ya no satisface ψ_r) y por lo tanto $\#\|\psi_r\|^{\mathcal{M}'} < \#\|\psi_{r-1}\|^{\mathcal{M}'}$, y también que $\|\psi_{r-1}\|^{\mathcal{M}} = \|\psi_r\|^{\mathcal{M}} = T$.
   - Si $\|\psi_{r-1}\|^{\mathcal{M}'} = S'_k$. Luego existe $\bar{h}' \in W' \setminus S'_k$ tal que para cada $\bar{h}_i \in W \setminus T$ hay un $\bar{\psi}_i^k$ donde $\mathcal{M}, \bar{h}_i \not\models \bar{\psi}_i^k$ pero $\mathcal{M}', \bar{h}' \models \bar{\psi}_i^k$.
     Sea $\varphi^k = \bigwedge_{i=0}^{m} \bar{\psi}_i^k$, $\psi_r = \psi_{r-1} \vee \varphi^k$:
     - Si $\|\psi_r\|^{\mathcal{M}'} \notin N'_{w'}$, entonces $\mathcal{M}, w \models \boxminus\psi_r$ pero $\mathcal{M}', w' \not\models \boxminus\psi_r$. Absurdo.
     - Si no, sea $\eta = \mathsf{E}(\varphi^k \wedge \neg\psi_{r-1})$, entonces $\mathcal{M}', w' \models \eta$ pero $\mathcal{M}, w \not\models \eta$. Absurdo.

3. La tercera condición es probada por analogía con la segunda condición.

4. Probemos ahora que la la relación resultante es *total*. Supongamos *ad absurdum* que no lo es. Sin pérdida de generalidad, supongamos que existe $v \in W$ tal que no es bisimilar a ningún elemento de W'. Sea $W' = \{w'_0, \ldots, w'_p\}$. Por nuestra definición de bisimulación,

esto significa que v no es equivalente a *ningún* mundo en W'. Entonces, para cada w'_i existe ψ_i tal que $\mathcal{M}, v \models \psi_i$ pero $\mathcal{M}', w'_i \not\models \psi_i$. Sea $\psi = \bigwedge_{i=0}^{p} \psi_i$. Entonces $\mathcal{M}, w \models E\psi$ pero $\mathcal{M}', w' \not\models E\psi$. Absurdo, dado que w y w' son modalmente equivalentes.

□

Nuevamente, podemos ver que la condición de modelo finito como hipótesis de la proposición 3.13 es necesaria usando los mismos contraejemplos que los empleados para $\mathcal{N}_=$ en los lemas 3.3 y 3.4.

Definición 3.12. Sean $\mathcal{M} = \langle W, N, P \rangle$, $\mathcal{M}' = \langle W', N', P' \rangle$ dos modelos minimales. Definimos una *bisimulación minimal strict total simétrica* entre $\mathcal{M}$ y $\mathcal{M}'$ como una bisimulación minimal strict total simétrica entre $\mathcal{M}^d$ y $\mathcal{M}'^d$.

Proposición 3.14. Las proposiciones 3.12 y 3.13 son válidas en la clase de todos los modelos minimales bajo bisimulación minimal strict simétrica.

Demostración. Inmediato de la proposición 3.4 y la definición 3.12. □

4. AXIOMATIZACIÓN Y COMPLEJIDAD

En este capítulo partiremos de los resultados del capítulo 3 para dar un análisis de $\mathcal{N}_=$ y $\mathcal{N}_\subseteq$ en términos axiomáticos y de complejidad.

4.1 Sistema axiomático

Como ya mencionamos, tanto *Neighbourhood Semantics Strict* como *Neighbourhood Semantics Loose* no fuerzan omnisciencia lógica, que podría expresarse como "siempre que un agente conoce todas las fórmulas de un conjunto Ψ, y Ψ implica lógicamente la fórmula φ, entonces el agente también conoce φ". En la lógica modal básica el axioma K y la regla de generalización garantizan omnisciencia lógica. Mientras que en $\mathcal{N}_=$ y $\mathcal{N}_\subseteq$ esta forma fuerte de omnisciencia no se cumple (ver el contraejemplo de la figura 2.1 en la sección 2.3).

Además de esta noción de omnisciencia lógica, hay otras nociones más débiles que la anterior, que también son interesantes para la lógica epistémica. Veremos algunas de ellas, caracterizadas por los siguientes esquemas que se pueden encontrar en [3]:

N. $\Box\top$

C. $\Box A \wedge \Box B \rightarrow \Box(A \wedge B)$

M. $\Box(A \wedge B) \rightarrow \Box A \wedge \Box B$

Cada uno de estos esquemas tiene un significado epistémico. **M** expresa que los agentes pueden realizar inferencias de hechos que conocen, **C** que los agentes pueden combinar hechos, y **N** que los agentes creen en hechos evidentemente verdaderos, es decir, en verdades universales. Cumplen una función importante dado que nos dan una idea de la clase de poder de razonamiento que los agentes pueden tener en lógicas más débiles que **K**.

Veremos que, si bien estos esquemas son válidos en cualquier clase de modelos relacionales, tienen contraejemplos en modelos minimales bajo $\mathcal{N}_=$ y $\mathcal{N}_\subseteq$.

Proposición 4.1. [3, Teorema 7.4] Ninguno de los esquemas **N** y **C** es válido en la clase de todos los modelos, bajo las semánticas $\mathcal{N}_=$ y $\mathcal{N}_\subseteq$. Además, **M** tampoco es válido en $\mathcal{N}_=$.

Demostración. Para cada esquema es suficiente describir una instancia y un modelo minimal que la falsifica en ambas semánticas. Como hicimos antes, escribiremos $\Box$ cuando no nos interesa diferenciar entre $\boxminus$ y $\boxdot$.

Para **N**. Aquí la instancia es simplemente $\Box\top$. Sea $\mathcal{M} = \langle W, N, P\rangle$ un modelo minimal de forma tal que $W = \{w\}$ y $N_w = \emptyset$ (sin importar qué sea P). Así, $\mathcal{M}$ contiene solamente un mundo w, y en ese mundo no hay proposiciones necesarias en $\mathcal{N}_=$ ni $\mathcal{N}_\subseteq$. En particular, la proposición W –i.e. $\|\top\|^{\mathcal{M}}$– no es necesaria en w, por lo tanto $\Box\top$ es falso en w para ambas semánticas.

Para **M**. Consideremos la instancia $\boxminus(p_0 \wedge p_1) \to (\boxminus p_0 \wedge \boxminus p_1)$, y sea $\mathcal{M} = \langle W, N, P\rangle$ un modelo minimal tal que $W = \{w, v\}$ (distintos), $N_w = \{\emptyset\}$, $P(p_0) = \{w\}$, $P(p_1) = \{v\}$ como en la figura 4.1. Entonces ni la proposición $\{w\}$ expresada por p_0 ni la proposición $\{v\}$ expresada por p_1 son necesarias en w bajo $\mathcal{N}_=$. Entonces $\boxminus(p_0 \wedge p_1)$ es verdadero en w, mientras que $\boxminus p_0$ y $\boxminus p_1$ son ambas falsas en w.

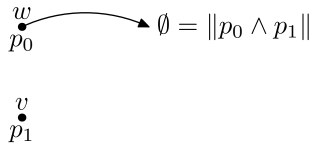

Fig. 4.1: $\mathcal{N}_=$ no cumple el esquema **M**.

Para **C**. Consideremos la instancia $(\Box p_0 \wedge \Box p_1) \to \Box(p_0 \wedge p_1)$, y sea $\mathcal{M}$ un modelo minimal como el de arriba excepto que $N_w = \{\{w\}, \{v\}\}$. Entonces, la situación de arriba es la reversa: las proposiciones expresadas por p_0 y p_1 son ambas necesarias en w, mientras que la proposición expresada por la conjunción no es necesaria en w. Así $\Box p_0$ y $\Box p_1$ son verdaderos en w bajo ambas semánticas, y $\Box(p_0 \wedge p_1)$ es falso en w tanto para $\mathcal{N}_=$ como para $\mathcal{N}_\subseteq$. De esto se sigue que $(\Box p_0 \wedge p_1) \to \Box(p_0 \wedge p_1)$ es falso en w. $\Box$

De hecho, la lógica $\mathcal{N}_=$ no sólo falsifica los esquemas **M**, **C** y **N**, sino que lo único que cumple es *clausura por equivalencia lógica*: "si el agente conoce φ, y φ y ψ son lógicamente equivalentes, entonces el agente conoce ψ". Esto está dado por:

$$\text{Si } p \leftrightarrow q \text{ es teorema, entonces } \boxminus p \leftrightarrow \boxminus q \text{ es teorema.}$$

Esta regla de inferencia, junto con *Modus Ponens*, *Substitución Uniforme*, y los teoremas de la lógica proposicional, forman una axiomatización completa y correcta de la lógica $\mathcal{N}_=$ [4, Teorema 9.2.2]. Si además tenemos el esquema **M** como axioma, será una axiomatización para $\mathcal{N}_\subseteq$. Esto se debe a que, como veremos, la inclusión de **M** como axioma restringe los modelos a aquellos que tienen una relación de neighbourhood cerrado bajo superconjuntos. En estos modelos, la validez de fórmulas en $\mathcal{N}_=$ coincide con la validez de fórmulas en la clase de todos los modelos de $\mathcal{N}_\subseteq$ (esto se verá en la proposición 4.3).

Cada uno de los esquemas vistos tiene asociado una condición a la relación N de los modelos minimales que define la clase de modelos en las cuales cada esquema es válido:

(m) si $X \subseteq Y$ y $X \in N_\alpha$, entonces $Y \in N_\alpha$

(c) si $X \in N_\alpha$ y $Y \in N_\alpha$, entonces $X \cap Y \in N_\alpha$

(n) $W \in N_\alpha$

Vale la pena mencionar que si bien no es evidente que la condición (m) caracterice la clase de modelos que validan **M**, se puede ver que la condición (m') sí lo hace, y que la condición (m) y (m') son equivalentes.

(m') si $X \cap Y \in N_\alpha$, entonces $X \in N_\alpha$ y $Y \in N_\alpha$

Lema 4.1. Los esquemas (m) y (m') son equivalentes.

Demostración. Supongamos que vale (m), y que X e Y son proposiciones en $\mathcal{M}$ tal que $X \cap Y \in N_a$. Entonces $Y \in N_a$, por (m), dado que $X \cap Y \subseteq Y$. Para la otra dirección, supongamos que vale (m'), y que X e Y son proposiciones en $\mathcal{M}$ tales que $X \subseteq Y$ y $X \in N_a$. Entonces $X = X \cap Y$, así que N_a contiene $X \cap Y$ y por lo tanto, por (m'), Y. □

Las condiciones (m), (c) y (n) definen las clases de modelos en donde sus respectivos esquemas se validan [3, Teorema 7.5], tanto en $\mathcal{N}_=$ como $\mathcal{N}_\subseteq$. Pero debemos observar que en $\mathcal{N}_\subseteq$, la condición (m) no modificará la satisfacción de ninguna fórmula en la lógica. Esto se debe a que esta relación de clausura por superconjuntos se encuentra expresada en la relación de satisfacibilidad de $\mathcal{N}_\subseteq$ y es por ello que **M** es válida en la clase de todos los modelos.

Proposición 4.2. El esquema **M** es válido en la clase de todos los modelos minimales, bajo Neighbourhood Semantics Loose.

Demostración. Si $\mathcal{M}, w \models \boxtimes(A \wedge B)$, entonces hay un elemento $X \in N_w$ tal que $X \subseteq \|A \wedge B\|^{\mathcal{M}}$, pero $\|A \wedge B\|^{\mathcal{M}} = \|A\|^{\mathcal{M}} \cap \|B\|^{\mathcal{M}} \subseteq \|A\|^{\mathcal{M}}$ como lo muestra la figura 4.2, por lo que $X \subseteq \|A\|^{\mathcal{M}}$. Esto significa que $\mathcal{M}, w \models \boxtimes A$. De igual forma operamos para probar que $\mathcal{M}, w \models \boxtimes B$.

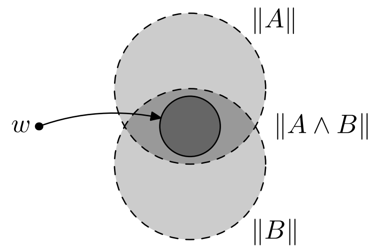

Fig. 4.2: $\mathcal{N}_\subseteq$ cumple el esquema **M**.

□

Esto significa que $\mathcal{N}_\subseteq$ es una lógica menos general en el sentido de que "el agente puede hacer inferencias". Es decir que si el agente conoce $A \wedge B$ entonces conoce A. Epistemológicamente diremos que tiene un mayor poder de razonamiento. Mientras que en $\mathcal{N}_=$ el agente no tiene por qué poder hacer este tipo de inferencias, o en todo caso para poder hacerlas deberá restringirse a un modelo condicionado por (m) en donde se valide **M**.

4.2 Complejidad de $\mathcal{N}_\subseteq$

Por lo visto antes, y la dualidad entre satisfacibilidad y validez, claramente las nociones de satisfacción en $\mathcal{N}_=$ y $\mathcal{N}_\subseteq$ no son la misma. Nuestro aporte original, consistirá en establecer la complejidad del problema satisfacibilidad para $\mathcal{N}_\subseteq$, relacionando ambas lógicas para poder transferir resultados conocidos de $\mathcal{N}_=$.

Definiremos a continuación la clase de modelos minimales en donde el esquema **M** es válido para $\mathcal{N}_=$.

Definición 4.1 (Suplementación [3])**.** Sea $\mathcal{M} = \langle W, N, P\rangle$ un modelo minimal. La *suplementación* de $\mathcal{M}$ es el modelo minimal $\mathcal{M}^+ = \langle W, N^+, P\rangle$ en dónde N_a^+ es la clausura por superconjuntos de N_a, para cada a en $\mathcal{M}$. Es decir, para cada a y X en $\mathcal{M}$,

$$X \in N_a^+ \iff Y \subseteq X \text{ para algún } Y \in N_a.$$

De hecho, la diferencia entre el poder de razonamiento de $\mathcal{N}_=$ y $\mathcal{N}_\subseteq$ está exactamente definido por el esquema **M**. Mostraremos que la noción de satisfacibilidad de $\mathcal{N}_\subseteq$ coincide con la de $\mathcal{N}_=$ restringido a la clase de modelos que validan **M** (esto es, la clase de modelos suplementados).

Proposición 4.3. Como ya hemos mencionado, denotaremos $\models_\subseteq$ a la relación de satisfacibilidad para $\mathcal{N}_\subseteq$ y $\models_=$ a la de $\mathcal{N}_=$. Consideremos además fórmulas donde el operador modal $\Box$ será interpretado según la semántica indicada en cada caso. Entonces, para cada fórmula modal φ y modelo minimal $\mathcal{M} = \langle W, N, P\rangle$, $\forall w \in W$:

$$\mathcal{M}, w \models_\subseteq \varphi \iff \mathcal{M}^+, w \models_= \varphi$$

o de forma equivalente:

$$\|\varphi\|_\subseteq^{\mathcal{M}} = \|\varphi\|_=^{\mathcal{M}^+}$$

donde $\mathcal{M}^+ = \langle W, N^+, P\rangle$ es la suplementación de $\mathcal{M}$.

Demostración. Por inducción en φ:

<u>Caso base</u>: Si φ es una variable proposicional, podemos probar trivialmente que $\mathcal{M}, w \models_\subseteq \varphi \iff \mathcal{M}^+, w \models_= \varphi$ dado que $\mathcal{M}$ y $\mathcal{M}'$ tienen la misma función de valuación P.

<u>Paso inductivo 1</u>: Si $\varphi = \psi_1 \wedge \psi_2$, por hipótesis inductiva sabemos que $\mathcal{M}, w \models_\subseteq \psi_1 \iff \mathcal{M}^+, w \models_= \psi_1$ y $\mathcal{M}, w \models_\subseteq \psi_2 \iff \mathcal{M}^+, w \models_= \psi_2$. Por lo tanto, $\mathcal{M}, w \models_\subseteq \psi_1$ y $\mathcal{M}, w \models_\subseteq \psi_2$ si y sólo si $\mathcal{M}^+, w \models_= \psi_1$ y $\mathcal{M}^+, w \models_= \psi_2$.

Si $\varphi = \neg\psi$, por hipótesis inductiva sabemos que $\mathcal{M}, w \models_\subseteq \psi \iff \mathcal{M}^+, w \models_= \psi$ entonces $\mathcal{M}, w \not\models_\subseteq \psi \iff \mathcal{M}^+, w \not\models_= \psi$, esto es, $\mathcal{M}, w \models_\subseteq \neg\psi \iff \mathcal{M}^+, w \models_= \neg\psi$.

<u>Paso inductivo 2</u>: Supongamos que $\varphi = \Box\psi$

$\Rightarrow$) Por hipótesis $\mathcal{M}, w \models_\subseteq \Box\psi$, esto significa $\exists X \in N_w$ tal que $X \subseteq \|\psi\|_\subseteq^{\mathcal{M}}$. Por definición de suplementación (+) esto implica que $\|\psi\|_\subseteq^{\mathcal{M}} \in N_w^+$. Aplicando la hipótesis inductiva, $\|\psi\|_\subseteq^{\mathcal{M}} = \|\psi\|_=^{\mathcal{M}^+}$, esto significa que $\|\psi\|_=^{\mathcal{M}^+} \in N_w^+$. Luego, $\mathcal{M}^+, w \models_= \Box\psi$

$\Leftarrow$) Por hipótesis $\mathcal{M}^+, w \models_= \Box\psi$, esto es $\|\psi\|_=^{\mathcal{M}^+} \in N_w^+$. Por definición de suplementación, $\exists X \in N_w$ tal que $X \subseteq \|\psi\|_=^{\mathcal{M}^+}$. Pero por hipótesis inductiva $\|\psi\|_=^{\mathcal{M}^+} = \|\psi\|_\subseteq^{\mathcal{M}}$, en otras palabras, $X \subseteq \|\psi\|_\subseteq^{\mathcal{M}}$ Esto significa que $\mathcal{M}, w \models_\subseteq \Box\psi$. $\Box$

Corolario 4.1. Sea $\mathcal{S}$ la clase de modelos suplementados. Para cada fórmula modal $\varphi \in \mathcal{N}_\subseteq$ y modelo minimal $\mathcal{M} \in \mathcal{S}$:

$$\mathcal{S} \models_\subseteq \varphi \iff \mathcal{S} \models_= \varphi.$$

Demostración. Cada $\mathcal{M} \in \mathcal{S}$ verifica que $\mathcal{M}^+ = \mathcal{M}$ por idempotencia de la suplementación. Sólo resta aplicar la proposición 4.3. $\Box$

Proposición 4.4. El problema de satisfacción en $\mathcal{N}_\subseteq$ está en *NP*.

Demostración. En [14] se demuestra que el problema de satisfacción en $\mathcal{N}_=$ restringido a la clase $\mathcal{S}$ (la clase de modelos suplementados $\varepsilon_{\{3\}}$) está en *NP*.

Si queremos saber si φ es satisfacible en $\mathcal{N}_\subseteq$, podemos dar φ como input de la máquina de Turing *NP* que resuelve el problema de satisfacción para $\mathcal{N}_=$ en la clase de modelos suplementados.

- Si responde *sí*, entonces debe haber un modelo suplementado $\mathcal{M}^+$ tal que $\mathcal{M}^+, w \models_= \varphi$. Luego podemos instanciar la proposición 4.3 con $\mathcal{M}^+$ y, como $\mathcal{M}^{++} = \mathcal{M}^+$, concluir que $\mathcal{M}^+, w \models_\subseteq \varphi$. Luego, φ es satisfacible en $\mathcal{N}_\subseteq$.
- Si responde *no*, supongamos *ad absurdum* que $\exists \mathcal{M}$ tal que $\mathcal{M}, w \models_\subseteq \varphi$. Por la proposición 4.3 esto implica que $\mathcal{M}^+, w \models_= \varphi$. Pero como el algoritmo ha respondido *no*, no puede haber un modelo suplementado $\mathcal{M}^+$ que satisfaga φ bajo $\mathcal{N}_=$. Absurdo. Luego, φ es insatisfacible en $\mathcal{N}_\subseteq$.

□

4.3 Complejidad de $\mathcal{N}_=(\mathsf{E})$

Siendo $\mathcal{N}_=(\mathsf{E})$ una lógica interesante desde el punto de vista de teoría de modelos, daremos, como aporte original, una cota superior (*upper bound*) a la complejidad de SAT-$\mathcal{N}_=(\mathsf{E})$. Esto lo haremos utilizando un argumento basado en *conjuntos de Hintikka*. Para otros ejemplos de utilización de este tipo de técnica, referirse a [2] para la lógica PDL y [11] para la lógica modal básica con modalidad universal. Pero antes, introduciremos el concepto de conjunto de Hintikka adaptado a $\mathcal{N}_=(\mathsf{E})$.

Definición 4.2. Un conjunto de subfórmulas Σ se dice:

Cerrado bajo subfórmulas: Si $\sigma \in \Sigma$ y θ es una subfórmula de σ, entonces $\theta \in \Sigma$.

Cerrado bajo negaciones simples: Si $\sigma \in \Sigma$ y σ no es de la forma $\neg\theta$, entonces $\neg\sigma \in \Sigma$.

Definición 4.3. Sea Σ un conjunto de fórmulas en $\mathcal{N}_=(\mathsf{E})$, y sea $\neg\Sigma$ la clausura bajo subfórmulas y negaciones simples de Σ. Un *conjunto de Hintikka* H sobre Σ es cualquier subconjunto maximal de $\neg\Sigma$ que satisface las siguiente condiciones:

(i) si $\neg\varphi \in \neg\Sigma$, entonces $\neg\varphi \in H$ sii $\varphi \notin H$.

(ii) si $\varphi \wedge \psi \in \neg\Sigma$, entonces $\varphi \wedge \psi \in H$ sii $\varphi \in H$ y $\psi \in H$.

(iii) si $\mathsf{E}\varphi \in \neg\Sigma$ y $\varphi \in H$, entonces $\mathsf{E}\varphi \in H$

Denotaremos al conjunto de todos los conjuntos de Hintikka sobre Σ con $Hin(\Sigma)$.

Proposición 4.5. El problema de satisfacción de $\mathcal{N}_=(\mathsf{E})$ está en *EXP-Time*

Demostración. Por construcción a partir de conjuntos de Hintikka. Sea $\varphi \in \mathcal{N}_=(\mathsf{E})$ una fórmula arbitraria, y $\Sigma = \{\varphi\}$.

Sea $\Delta \subseteq \wp(Hin(\Sigma))$, que consiste en todos los $C_i \subseteq Hin(\Sigma)$ que satisfacen las siguientes propiedades:

(i) $\forall H, H' \in C_i, \forall \mathsf{E}\varphi \in \neg\Sigma : \mathsf{E}\varphi \in H \iff \mathsf{E}\varphi \in H'$,

(ii) $\forall H \in C_i, \forall \mathsf{E}\varphi \in \neg\Sigma$: si $\mathsf{E}\varphi \in H$ entonces $\exists H' \in C_i$ tal que $\varphi \in H'$,

(iii) No existe $D \subseteq Hin(\Sigma)$, $C_i \subsetneq D$, que satisfaga (i) y (ii).

Como $\neg\Sigma$ Contiene a lo sumo $2|\varphi|$ elementos, deben existir a lo sumo $2^{2|\varphi|}$ conjuntos maximales C_i que cumplen esta condición. Entonces, el tamaño de $\Delta = \{C_1, \ldots, C_n\}$ es a lo sumo exponencial en el tamaño de φ.

Dado cualquier $W \subseteq \wp(\mathcal{N}_{=}(\mathsf{E}))$, podemos definir $[\varphi] = \{X \mid X \in W \wedge \varphi \in X\}$.

Para cada $C_i \in \Delta$ construimos una estructura $\mathcal{M}_i = \langle W, N, V\rangle$. Definimos $W := C_i$. Definamos la relación neighbourhood N asociado a cada $H \in W$ de la siguiente manera:

$$\boxminus\varphi \in H \text{ sii } [\varphi] \in N_H.$$

Finalmente definimos V con $V(p) = \{H \in W \mid p \in H\}$, para todas las variables proposicionales $p \in \mathsf{PROP}$.

Dado $\psi \in \mathcal{N}_{=}(\mathsf{E})$, comprobaremos su satisfacibilidad como sigue. Definiendo $\Sigma = \{\psi\}$, construiremos $Hin(\Sigma)$ y ejecutaremos el proceso recién descrito. Este proceso termina devolviendo un conjunto de estructuras $\{\mathcal{M}_1, \ldots, \mathcal{M}_n\}$. En breve probaremos la siguiente proposición, para toda fórmula $\varphi \in \neg\Sigma$:

$$\varphi \text{ es satisfacible sii } \varphi \in H \text{ para algún } H \in \mathcal{M}_i, 1 \leq i \leq n. \tag{4.1}$$

Si podemos probar esta afirmación, el teorema se sigue inmediatamente. Para ver esto, notemos que el número de conjuntos de Hintikka sobre Σ es a lo sumo exponencial en el tamaño de ψ. Luego el número de conjuntos de Hintikka en cada C_i es también a lo sumo exponencial. Como ya hemos visto antes, hay una cantidad a lo sumo exponencial de C_i, y por lo tanto el proceso es un algoritmo *EXP-Time*.

Así, sólo resta probar (4.1). Para la dirección de derecha a izquierda, mostraremos que si $\varphi \in H$ para algún $H \in \mathcal{M}_i$, entonces $\mathcal{M}_i$ en sí mismo satisface φ en H. De hecho mostraremos que para todo $\varphi \in \neg\Sigma$ y todo $H \in W_i$,

$$\mathcal{M}_i, H \models \varphi \text{ sii } \varphi \in H \tag{4.2}$$

Esta prueba es por inducción. La cláusula para símbolos proposicionales es simple, y el paso de combinaciones booleanas se sigue usando las cláusulas (i) y (ii) en la definición de conjuntos de Hintikka.

Para el paso que involucra operadores modales necesitamos probar la siguiente afirmación para todo $\mathcal{M}_i$:

$$\forall\boxminus\varphi \in \neg\Sigma, \boxminus\varphi \in H \text{ sii } [\varphi] \in N_H \tag{4.3}$$

Para comprobar su validez, simplemente debemos notar que (4.3) es exactamente la definición que hemos dado para la función de neighbourhoods N al construir cada $\mathcal{M}_i$.

Finalmente, para probar (4.2) en el paso que involucra la modalidad universal, supongamos que $\mathcal{M}_i, H \models \mathsf{E}\varphi$ entonces, por hipótesis inductiva, $\varphi \in H'$ para algún $H' \in W$. Supongamos que $\mathsf{E}\varphi \notin H$, entonces $\mathsf{E}\varphi \notin H''$ para cada H'' en $\mathcal{M}_i$, por condición (i) de la definición de Δ. Por el contrarecíproco de la condición (iii) de la definición de conjuntos de Hintikka, $\varphi \notin H''$ para cada H'', incluyendo por supuesto a H', lo que contradice nuestra primera presunción. Para demostrar la otra dirección, supongamos que $\mathsf{E}\varphi \in H$. Entonces, por la condición (ii) impuesta al definir Δ debe haber al menos un H' en donde $\varphi \in H'$. Por hipótesis inductiva, $\mathcal{M}_i, H' \models \varphi$ y por consiguiente $\mathcal{M}_i, H \models \mathsf{E}\varphi$.

Para completar (4.1) basta mostrar que si φ es satisfacible, entonces es satisfacible en algún $\mathcal{M}_i$.

Supongamos que φ es satisfacible en algún modelo $\mathcal{M}$. Sea $S_{\mathcal{M}}$ el conjunto de subconjuntos de $\neg\Sigma$ que efectivamente ocurren en $\mathcal{M}$, es decir, $S_{\mathcal{M}} = \{H \in Hin(\Sigma) \mid \mathcal{M}, w \models H \text{ para algún } w \in \mathcal{M}\}$. Obviamente, $S_{\mathcal{M}} \subseteq Hin(\Sigma)$, y cada elemento de $S_{\mathcal{M}}$ contiene las mismas fórmulas del tipo $\mathsf{E}\varphi$. Entonces debe existir un conjunto $C_i \in \Delta$ tal que $S_{\mathcal{M}} \subseteq C_i$. Así, $\varphi \in H$, para algún $H \in C_i$. $\square$

5. CONCLUSIONES, PREGUNTAS ABIERTAS Y TRABAJO A FUTURO

Conclusión

A lo largo de este trabajo hemos realizado una comparación de expresividad entre las dos lógicas $\mathcal{N}_=$ y $\mathcal{N}_\subseteq$.

Por un lado a través de su noción de bisimulación. Vimos que mientras en $\mathcal{N}_\subseteq$ la noción de bisimulación coincide con la intuición de la definición de la semántica, en el caso de $\mathcal{N}_=$ la bisimulación resulta compleja. Analizamos la definición 3.7 intuitiva y simétrica de la bisimulación para $\mathcal{N}_\subseteq$ pero resulta ser demasiado estricta: $\mathcal{N}_=$ no puede diferenciar superconjuntos de neighbourhoods, algo que la bisimulación simétrica sí logra discriminar. Investigamos cuál es el poder expresivo que necesitaríamos agregarle a $\mathcal{N}_=$ para llegar a aquella noción intuitiva de bisimulación. Finalmente encontramos que $\mathcal{N}_=(\mathsf{E})$ incluye justamente este poder y en consecuencia dimos su noción de bisimulación.

Por otro lado, analizamos la expresividad de estas lógicas a través de la complejidad computacional del problema de satisfacción. Dimos un argumento para vincular ambas lógicas usando la noción de *suplementación* en los modelos y a partir de esto demostramos que SAT-$\mathcal{N}_\subseteq$ está en *NP*. Dado que $\mathcal{N}_=(\mathsf{E})$ resultó ser una lógica interesante desde la perspectiva de bisimulación, es de interés conocer su complejidad computacional. Utilizando un argumento de selección adaptado a $\mathcal{N}_=(\mathsf{E})$ basado en conjuntos de Hintikka, probamos que SAT-$\mathcal{N}_=(\mathsf{E})$ está en *EXP-Time*.

Próximos pasos

Caracterización. Hemos visto en la sección 1.4.1 que para la semántica relacional, la noción de bisimulación nos permite identificar las fórmulas de la lógica de primer orden correspondientes al lenguaje modal básico. Para llegar a un resultado como este en *Neighbourhood Semantics*, es necesario tener una traducción estándar de las fórmulas de cada semántica a una lógica más estudiada y conocida.

- Como en las lógicas de *Neighbourhood Semantics* los conjuntos de mundos son ciudadanos de primer orden, posiblemente no podremos hacer la traducción a fórmulas de primer orden como en la semántica relacional, seguramente sí a lógica de segundo orden. Pero ¿podemos encontrar una traducción estándar intuitiva a una lógica más débil que segundo orden? (Por ejemplo a una lógica de primer orden con *sorts*.)

- Una vez encontrada una traducción estándar para $\mathcal{N}_=$ ó $\mathcal{N}_\subseteq$, ¿se podrá probar algún resultado de caracterización similar al resultado que van Benthem demostró para la lógica modal básica?

Lenguaje infinito en $\mathcal{N}_=$. Hemos visto que tanto para el teorema 1.2 de Hennessy-Milner de la lógica modal básica como para la proposición 3.2 para $\mathcal{N}_\subseteq$, la hipótesis de finitud en cada caso, está directamente relacionada con la finitud de las fórmulas de la lógica. Sin embargo, la demostración de la proposición 3.6 para $\mathcal{N}_=$ no nos permite debilitar la hipótesis de modelo finito con la inclusión de conjunciones infinitas en nuestras fórmulas. Esto nos lleva a preguntarnos:

- En un lenguaje modal con conjunciones infinitas, ¿será verdadera la proposición 3.6 para modelos infinitos?
- De ser así, ¿es necesario alguna hipótesis de finitud (neighbourhood-finito, co-neighbourhood-finito, imagen-finita)?

Cota inferior para $\mathcal{N}_=(\mathsf{E})$. Hemos visto con la proposición 4.5 que el problema SAT-$\mathcal{N}_=(\mathsf{E})$ está en *EXP-Time*. Pero, ¿cuál es la la cota inferior (*lower bound*) para SAT-$\mathcal{N}_=(\mathsf{E})$? Como SAT-$\mathcal{N}_=$ está en *NP*, la cota inferior de $\mathcal{N}_=(\mathsf{E})$ debe estar entre *NP* y *EXP-Time*.

Es posible que esté en *EXP-Time*, porque en el caso de la lógica modal básica la inclusión de la modalidad universal cambia la complejidad de SAT-**K** de *P-Space* a *EXP-Time*. En este caso, se podría intentar encontrar una traducción t tal que para el lenguaje modal básico con modalidad universal,

$$\models_{\mathbf{K}} \varphi \text{ sii } \models_{=} t(\varphi).$$

Posiblemente esta traducción incluirá instancias de los esquemas **N**, **C** y **M** para llevar el modelo minimal a un modelo de cierto tipo[1] que es equivalente a un modelo relacional [3, teorema 7.9].

Sin embargo también hay razones para pensar que el problema podría estar en *NP*. Teniendo en cuenta que en la clase de todos los modelos la modalidad universal por sí misma (sin ningún otro operador modal) está en *NP*. Es justamente la interacción entre $\Diamond$ y E que lleva SAT-$\mathbf{K}(\mathsf{E})$ a *EXP-Time*. Sin embargo $\boxminus$ es un operador muy débil, y es posible que no tenga el suficiente poder expresivo para que se produzca el salto de *NP* a *EXP-Time* en SAT-$\mathcal{N}_=(\mathsf{E})$.

[1] Modelo *quasi-filter* o *aumentado* [3, definición 7.8 y sección 7.3].

BIBLIOGRAFíA

[1] Marco Aiello and Johan van Benthem. Logical patterns in space. In D. Barker-Plummer, D. Beaver, J. van Benthem, and P. Scotto di Luzio, editors, *Words, Proofs, and Diagrams*, pages 5–25. CSLI Stanford, 2002.

[2] Patrick Blackburn, Maarten de Rijke, and Yde Venema. *Modal Logic*, chapter 6. Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, Cambridge, UK, 2001.

[3] Brian F. Chellas. *Modal Logic. An Introduction*, chapter 7. Cambridge University Press, Cambridge, UK, 1980.

[4] Ronald Fagin, Joseph Y. Halpern, Yoram Moses, and Moshe Y. Vardi. *Reasoning About Knowledge*. The MIT Press, 1995.

[5] Jelle Gerbrandy. *Bisimulations on Planet Kripke*, chapter 2 and 3. ILLC Dissertation Series. ILLC, Universiteit van Amsterdam, Amsterdam, Netherlands, 1998.

[6] David Harel. Dynamic logic. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic: Volume II: Extensions of Classical Logic*, pages 497–604. Reidel, Dordrecht, 1984.

[7] Clarence Irving Lewis. *A Survey of Symbolic Logic*. University of California Press, Berkeley, 1918. Reprint of Chapters I–IV by Dover Publications, 1960, New York.

[8] Richard Montague. Pragmatics. In *Contemporary Philosophy: a Survey*, pages 102–122. La Nuova Italia Editrice, Florence, 1968.

[9] Richard Montague. Universal grammar. *Theoria*, 36:373–398, 1970.

[10] Dana Scott. Philosophical problems in logic. In *Advice on modal logic*, pages 143–173. Reidel, 1968.

[11] Edith Spaan. *Complexity of Modal Logics*, chapter 2. Haveka B.V., Alblasserdam, The Netherlands, 1993.

[12] Johan van Benthem. *Modal Correspondence Theory*. PhD thesis, Mathematisch Instituut & Instituut voor Grondslagenonderzoek, University of Amsterdam, 1976.

[13] Johan van Benthem. Dynamic bits and pieces. In *Exploring Logical Dynamics*. CSLI Publications, Stanford, 1997. Annual update.

[14] Moshe Y. Vardi. On the complexity of epistemic reasoning. In *Proceedings 4th IEEE Symposium on Logic in Computer Science*, pages 243–252, 1989.