FAMAF Facultad de Matemática, Astronomía, Física y Computación

UNC Universidad Nacional de Córdoba

Lógicas epistémicas basadas en habilidades

por

Andrés Román Saravia

Presentado ante la Facultad de Matemática, Astronomía, Física y Computación como parte de los requerimientos para la obtención del grado de Doctor en Ciencias de la Computación de la

UNIVERSIDAD NACIONAL DE CÓRDOBA

Junio, 2024

Director: Raul Fervari

Tribunal Especial (titulares y suplentes):
Dr. Pedro D'Argenio (FAMAF-UNC, CONICET)
Dr. Miguel Campercholi (FAMAF-UNC, CONICET)
Dr. Santiago Figueira (UBA, CONICET)
Dr. Nazareno Aguirre (UNRC, CONICET)
Dr. Sergio Abriola (UBA, CONICET)

A todos los que me preceden y a todos los que me sucedan.
Y en especial a todos los que están.

Agradecimientos

Love is wise — hatred is foolish.

Bertrand Russell

Estos últimos cuatro años han sido los más complicados pero más transformativos de mi vida. En el medio ha ocurrido una pandemia mundial, he trabajado en proyectos importantes, he viajado bastante y he aprendido muchas cosas tanto adentro como fuera del doctorado. Durante una buena parte del mismo, especialmente a la hora de escribir este trabajo, mi mente no ha parado de llenarse de pensamientos sobre el pasado, el presente y el futuro. A veces no puedo dejar de reflexionar que la vida cambia tan rápidamente, que difícilmente nos da tiempo para adaptarnos a los escenarios, y lo único que nos queda es aceptarlos, adaptarnos y seguir sin más. O peor aún, nos da el tiempo y los recursos pero nuestra visión sesgada es tan grave, que decidimos ir por el autosaboteo. La incertidumbre, como en este trabajo, siempre será una constante en la vida. Más aún considerando el contexto político actual.

En el pasado, el comienzo de mi tesis de licenciatura en ciencias de la computación rezaba "a todos los que me preceden y a todos los que me sucedan", en un intento de honrar a la gente que me antecede en esta área y motivar a los que me sucedan en la misma. En este caso, después de cuatro años, decidí darle una vuelta y agregar "y en especial a todos los que están". Esta tesis se la dedico a todos aquellos que me han acompañado en este trayecto, independientemente si siguen estando conmigo o no ahora, o en un futuro.

En primer lugar, quiero agradecer a mi director de tesis Raul Fervari. Desde 2018, ha sido una piedra fundacional en mi pensamiento científico más maduro y una gran inspiración de mi actitud de investigador. Agradezco haber tenido la fortuna de contar con él a lo largo de este tedioso pero gratificante proceso y sólo deseo lo mejor para él. Quizás de lo que me lamento es no haber tenido más conversaciones de las que ya tuvimos. Y, por supuesto, de las cantidades ingentes de porquerías que escribí en el borrador de esta tesis que habrá tenido que leer. Estoy seguro que más de una vez se habrá llevado las manos a la cabeza leyéndolas. Si pudiera, le pagaría la tinta invisible de la lapicera digital que ha tenido que usar.

También quiero agradecer a las personas con las que trabajé en estos años: Carlos Areces, Valentin Cassano, Pablo Castro y Fernando Velázquez-Quesada. Y a gente con la quien no he podido trabajar pero que afortunadamente me he encontrado a lo largo de este camino: Luis Felipe Bartolo Alegre, Seungrak Choi, Mina Young Pedersen, John Elias Lindqvist, Rustam Galimullin y muchos más. Agradezco muchísimo el haber sido introducido al mundo de la investigación de la mano de todos ellos y ojalá pueda seguir coincidiendo con ellos y con más personas.

Quiero agradecer a la Universidad Nacional de Córdoba y FaMAF, por darme la educación y el espacio para desarrollarme social y académicamente. Y al CONICET, por darme la oportunidad de ser un becario doctoral y abrirme las puertas a un mundo distinto, uno que jamás hubiera imaginado que existía. Haber sido seleccionado de tantas solicitudes me llena de orgullo y me hace sentir que mi labor es reconocida, que es vista, que es validada, que importa. Si bien la situación actual es incierta, deseo fervientemente que estas instituciones sigan existiendo, mejorando y creando investigadores y profesionales, como siempre lo han hecho.

Terminando los agradecimientos, por así decirlo, académicos, quiero agradecer a Alexandra Elbakyan, fundadora de Sci-Hub, y a Library Genesis. El conocimiento libera. Y nadie debe ser excluido de él.

Por otro lado, quiero agradecer a toda mi familia, en especial mis padres por acompañarme en este largo trayecto, e inculcarme la paciencia y la perseverancia para superar la adversidad. Han sido, son y serán el faro de mi vida y el amor que me han sabido dar no será desperdiciado.

Gracias por aguantarme, aún en los momentos en los que hasta yo mismo no puedo hacerlo conmigo mismo. Hablando en general, quiero agradecer a mi familia que está en Argentina y también a la que está en el exterior, como en España y Estados Unidos, y que he tenido la suerte de conocer y hablar. Agradezco tanto las charlas que tuve con ellos en mis viajes, por más cortas que hayan sido.

Lo que sigue en sección son agradecimientos más específicos y demasiado personales. Son personas que han sido un empuje en mi vida en estos últimos años y siento que merecen una mención aunque sea. Principalmente, quiero agradecer a mis abuelas paterna y materna: Martha Elena Agüero y Lucrecia Pérez Costa. Si bien, no podrán verme con este título en mano, sé muy bien que donde quieran que estén, estoy seguro que estarán orgullosas de mí. En especial, Lucrecia ha sido una gran influencia en mi vida desde mi infancia, a tal punto que una o más frases de los inicios de los capítulos las saqué de un libro suyo. Las amo mucho y las extraño mucho. A Antonella Freijo que me ha acompañado desde antes de mi doctorado en 2019 hasta un tiempo después que empecé a escribirlo. Nunca podré terminar de agradecerte las alegrías y los momentos que me has dado, y los recuerdos que tengo los atesoraré por el resto de mi vida. A mis compañeros de la oficina 231/232, especialmente a Marina Palacio, Muriel Zampieri, Lucía Morey y Lupe Peñaranda. Si bien, mi asistencia fue muy irregular, agradezco mucho las charlas sobre nuestros doctorados y frustraciones varias. A Nicolás Hörmann y Alejandro Silva por las eternas tardes en las que me ocupaban la oficina y de paso merendábamos a base de mate y alfajores Fulbito. A Adrián Edelstein, Kim de Oro y Lucía Bustos, por las juntadas improvisadas tanto en Güemes como en noches de vicio en Discord. A Victoria Bigatton y a nuestras trasnoches aprendiendo programación. Algún día iré a Río Negro a visitarte. A Valentina Luna, Agostina Mosquera, Julieta Marioni y Paloma Ottonello, amistades de larguísima data, que a pesar del tiempo siguen estando presentes en cada momento. A Franja Morada, especialmente a Iván Bürcher, Anita, Matías, Nehuén, Marty, Ceci y Cele; por todos estos años en los que me he sentido bienvenido en la agrupación. Les deseo lo mejor. A los grupos de las mesitas de las LEF y de la sala de estudio que he conocido recién en 2023, que son muchísimos pero quiero destacar algunos: Leo, Mariana, Agustín, Mauro, Vicente (Pollo), Lautaro, Gonzalo, Pablo, los cuatro Francos que conozco (especialmente a Milana y a Diosques), Aaron, Dania, Ana, Paula, Sofía y seguro muchos más que me estoy olvidando. Aunque no se note, me han salvado en un momento complicado y estoy agradecido con ustedes. A Mariel Palacio, Rocío Fonseca, Sara Vegetti, Lucas Cardacci y Tomás Radiulovich y cada charla que tuvimos sobre análisis numérico, codornices, teoría de categorías, topología o simplemente la vida misma.

Podría seguir y seguir listando. Hay tantas personas con las que me he encontrado en mis viajes, en la facultad, en la vida. Pero el tiempo es limitado y las páginas también, por no decir que mi memoria es volátil. A veces es increíble ser consciente de que uno es parte de un todo, una amalgama de personas, experiencias y momentos. Ser la suma de cada aporte, de cada historia, de cada instante que pasamos con otros y nosotros mismos, e incluso ser mucho más que esa suma. El futuro parece ser más amplio que el que pude pensar en ese extraño enero de 2020, en donde sentía que había llegado tan lejos que no creía que algo nuevo me iba a sorprender. Más amplio que el que siquiera podía concebir en 2014, cuando estaba decidiendo qué carrera estudiar. Realmente, no sé qué me deparará a partir de ahora. Pero a veces uno siente que, con cada final, vuelve al principio. Sólo que con algo más. Una eterna vuelta al comienzo, enfrentándose a otro lienzo en blanco y a la incertidumbre que este presenta, pero con el ferviente deseo de experimentar con nuevas ideas.

Resumen

> El tiempo no se compone solamente de horas ni de minutos, sino de amor y de voluntad; tenemos poco tiempo cuando tenemos poco amor.
>
> Alexandre Vinetw

Las lógicas epistémicas son una rama de las lógicas modales enfocadas en la representación del conocimiento. En estas se modelan entidades autónomas, comúnmente llamadas agentes, que interactúan en un entorno, usualmente compartido, y poseen un grado de conocimiento sobre el mismo. Específicamente, las lógicas epistémicas de *saber cómo* estudian el conocimiento que tienen los agentes sobre sus propias habilidades. En general, un agente sabe cómo lograr un objetivo φ bajo una condición inicial ψ, si existe un curso de acción o plan tal que siempre puede ejecutarse bajo ψ, nunca falla y siempre conduce a situaciones donde vale φ.

En este trabajo presentamos una nueva semántica para una lógica epistémica de saber cómo. En este nuevo enfoque, los operadores de la lógica son interpretados sobre planes lineales en sistemas de transiciones etiquetadas (LTS), pero además se incluye una noción de indistinguibilidad epistémica entre planes para cada agente. La misma representa la percepción de los agentes y permite modelar entidades más "reales", por ejemplo, acerca de las limitaciones que cada agente posee con respecto a su conocimiento. Este enfoque resulta novedoso, ya que nos permite obtener una clara distinción entre la información óntica, los planes que están disponibles, y la información epistémica, los planes que el agente percibe como posibles.

Además de resultar una alternativa adecuada desde un punto de vista conceptual, la lógica obtenida posee propiedades interesantes, cuyo estudio resulta una parte importante de esta tesis. En primer lugar, introducimos una axiomatización correcta y fuertemente completa con respecto a la clase de todos los modelos. La misma refleja el comportamiento deseado en cuanto a las fórmulas válidas en esta semántica. Por otro lado, investigamos el comportamiento computacional de nuestra lógica. En particular, demostramos que el problema de decidir si una fórmula de esta lógica es satisfacible es NP-completo, mientras que el problema de model-checking está en P. Además, investigamos su poder expresivo mediante la definición de bisimulaciones, y su relación con propuestas anteriores.

Gracias a esta separación entre la información óntica y epistémica dada por la semántica presentada, es posible realizar un tratamiento de esta lógica, similar al utilizado en lógica epistémica clásica. En este sentido, nos centramos en la definición de operadores dinámicos que actualizan cada tipo de información, obteniendo así resultados de expresividad, decidibilidad y axiomatizaciones. Además, discutimos los desafíos de lidiar con este tipo de operadores y mostramos herramientas alternativas para resolver algunos de los obstáculos presentados, como por ejemplo extender la lógica base para obtener la expresividad necesaria en la definición de axiomas de reducción.

Finalmente, introducimos una interpretación deóntica de este tipo de operadores, un enfoque usual dentro de la lógica modal. La lógica definida tiene como objeto razonar acerca de las habilidades de los agentes, las normas con las que deben cumplir, y sobre cómo los agentes *cumplen conscientemente* o no con estas. Una vez más, utilizando las herramientas adquiridas en esta tesis, estudiamos axiomatizaciones y el comportamiento computacional de la lógica deóntica obtenida, como así también variantes de la misma.

Abstract

Dans ses écrits, un sage Italien
Dit que le mieux est l'ennemi du bien.

Voltaire

Epistemic Logics are a branch of Modal Logics that focused on the field of knowledge representation. These model autonomous entities, commonly called agents, that interact in an environment, usually shared, and have a degree of knowledge about it. Specifically, Epistemic Logics related to *knowing how* study the knowledge that agents have about their own abilities. In general, an agent knows how to achieve a goal φ under an initial condition ψ, if there exists a course of action or plan such that it can always be executed under ψ, never fails, and always leads to situations where φ is true.

In this work we present a new semantics for an epistemic logic of knowing how. In this new approach, the logic operators are interpreted on linear plans in Labeled Transition Systems (LTS), but a notion of epistemic indistinguishability between plans for each agent is also included. This notion represents the perception of the agents and allows modeling more "realistic" entities, for example, about the limitations that each agent has with respect to their knowledge. This approach is novel, since it allows us to obtain a clear distinction between ontic information, the plans that are available, and epistemic information, the plans that the agent perceives as possible.

In addition to being an adequate alternative from a conceptual point of view, the logic obtained has interesting properties, the study of which is an important part of this thesis. First, we introduce a correct and strongly complete axiomatization regarding the class of all models. This system reflects the desired behavior regarding the valid formulas in this semantics. On the other hand, we investigate the computational behavior of our logic. In particular, we show that the problem of deciding whether a formula of this logic is satisfiable is NP-complete, while the model-checking problem is P. Furthermore, we investigate its expressive power by defining bisimulations, and its relationship with previous proposals.

Due to this separation between the ontic and epistemic information given by the semantics presented, it is possible to carry out a treatment for this logic, similar to the one used in classical epistemic logic. In this direction, we focus on the definition of dynamic operators that update each type of information, thus obtaining results of expressiveness, decidability and axiomatizations. In addition, we discuss the challenges of dealing with this type of operators and show alternative tools to solve some of the obstacles presented, such as extending the base logic to obtain the expressiveness necessary in the definition of reduction axioms.

Finally, we introduce a deontic interpretation of this type of operators, a common approach within modal logic. The defined logic aims to reason about the abilities of agents, the rules with which they must comply, and about how agents *consciously comply* or not with these. Once again, using the tools acquired in this thesis, we study axiomatizations and the computational behavior of the deontic logic obtained, as well as variants of it.

Índice general

Parte I

Introducción

Capítulo 1

La lógica y las ciencias de la computación

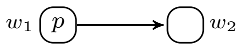
> Computer Science is a science of abstraction — creating the right model for a problem and devising the appropriate mechanizable techniques to solve it.
>
> Alfred Aho

1.1. Lógicas modales, ¿para qué?

En ciencias de la computación, un modelo es una representación abstracta de un sistema, un proceso, o más en general, de algún concepto u objeto específico relacionado a estos. Dicha representación tiene como objetivo simplificar, analizar o simular escenarios reales o teóricos vinculados a las ciencias de la computación, como así también comprender y resolver los problemas asociados a estos. En algunos casos, los mismos incluso nos ayudan a dar predicciones sobre determinados eventos. Por este motivo, existen situaciones en donde resulta necesario contar con herramientas que puedan captar de manera precisa las propiedades del objeto o situación en cuestión e incluso, utilizando mecanismos lógicos, obtener o inferir de manera rigurosa nueva información de estas representaciones o modelos. Desde el punto de vista computacional, el desafío suele ser encontrar los formalismos adecuados que cuenten con propiedades computacionales lo suficientemente buenas (por ejemplo, que sus tareas de razonamiento puedan efectuarse con baja complejidad computacional), con el objeto de automatizar o mecanizar su funcionamiento o sus procedimientos de inferencia. En este sentido, emerge una familia de lógicas que generalmente reúnen varios de estos requisitos: la familia de las *Lógicas Modales* [24, 26, 30, 105].

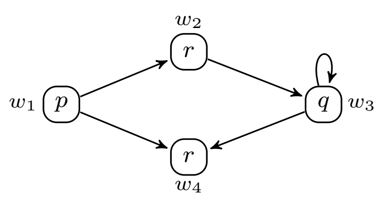
Informalmente, las lógicas modales son en general extensiones de la lógica proposicional, que cuentan con operadores capaces de describir "modos de verdad" o "modalidades" sobre determinadas propiedades. En términos generales, una modalidad es una palabra o expresión que se puede aplicar a una sentencia S, para crear una nueva que haga una aserción sobre el modo de verdad de S, sobre cuándo, dónde o cómo S es verdad, o bajo qué circunstancias S es verdadera [50]. Los orígenes de las lógicas modales pueden remontarse a Aristóteles [101], quien introduce distintas nociones de posibilidad y necesidad en sus silogismos. Por lo tanto, no solo razonamos acerca de si S es verdadera o falsa, si no que también podemos hacerlo acerca de si "es posible o necesario que S sea verdadera". Ejemplos de este tipo de sentencias son *"es posible que los lados de un rectángulo sean iguales"* o *"es necesario que la suma de los ángulos de un triángulo sea 180°"*. Además de necesidad y posibilidad, existen muchos otros conceptos que pueden ser caracterizados en el contexto de la lógica modal, tales como *temporalidad*, *conocimiento*, *creencias*, *obligaciones*, entre otras. Por este motivo no hablamos de una única lógica modal, sino de una familia de estas, formalizando los distintos conceptos que la situación nos requiera.

En la literatura moderna [24, 26, 113], las nociones de posibilidad y necesidad (así como otras nociones "modales") se formalizan mendiante los llamados "operadores modales" denotados como $\Diamond$ y $\Box$ respectivamente. Por ejemplo, dadas dos expresiones ψ y φ, la expresión modal $\Diamond\psi$ se lee como “es posible que ψ se cumpla”, mientras que la expersión $\Box\varphi$ se lee como “es necesario que φ se cumpla”. Semánticamente, las fórmulas modales se interpretan en la llamada “semántica de mundos posibles” en modelos relacionales [24,26,113]. La idea general de dicha semántica es que, a diferencia de la lógica proposicional donde existe una única interpretación de los hechos, existen diversas situaciones posibles donde cada hecho posee su propia interpretación. Por ejemplo, consideremos a p una variable representando el hecho *“hace frío”*. En lógica proposicional, esta situación se interpreta de manera *objetiva*: si hace frío, p se interpreta como verdadera, mientras que si no hace frío se interpreta como falsa. En lógica modal, esta situación se interpreta *subjetivamente:* si bien p puede ser verdadera actualmente, es *posible* que existan otras situaciones donde p es falsa, sin incurrir en una contradicción con el hecho actual. Por ejemplo, si bien puede hacer frío en la ubicación actual, es posible que no haga frío en otras ubicaciones remotas. En la semántica usual de mundos posibles [24, 26], esto se representa mediante modelos relacionales: un conjunto de estados o mundos etiquetados por los hechos o símbolos de proposición que son verdaderos en cada mundo posible, y una relación de accesibilidad estableciendo qué situaciones se consideran posibles a partir de un determinado mundo. Estos modelos en general también son llamados *Modelos de Kripke*. Luego, $\Diamond\psi$ es verdadera (en algún mundo w del modelo) si ψ es verdadera en *alguno* de los mundos considerados posibles a partir de w (es decir, en algún sucesor mediante la relación de accesibilidad). Por otro lado, $\Box\varphi$ es verdadera en w si φ es verdadera en *todos* los mundos posibles desde w. La situación descripta previamente puede visualizarse en el siguiente gráfico. En el mismo, se consideran dos ubicaciones posibles (w_1 y w_2). Si bien en w_1 objetivamente hace frío (p), desde el mismo se considera posible que no sea el caso en w_2 ($\Diamond p$).



Es evidente que los modelos mencionados anteriormente no son otra cosa que grafos dirigidos con etiquetas en los nodos. Los grafos son estructuras muy conocidas dentro de las ciencias de la computación, ya que nos permiten representar situaciones de las más diversas. Por ejemplo, es posible representar las relaciones dentro de una base de datos, el flujo de ejecución de un programa, las trayectorias o estrategias que un agente inteligente es capaz de ejecutar, o las conexiones dentro de una red social, por mencionar sólo algunas aplicaciones. Esta flexibilidad hace que las lógicas modales sean adecuadas para modelar y razonar acerca de tantas situaciones en ciencias de la computación. Además, la semántica de las lógicas modales es usualmente de “Primer Orden”, en el sentido de que los modelos relacionales son también modelos de Lógica de Primer Orden [37], por lo cual las lógicas modales pueden considerse como fragmentos de Primer Orden pero con un buen comportamiento computacional. Volveremos a los detalles de este hecho más tarde en esta tesis. A continuación presentamos un ejemplo que da cuenta de los detalles mencionados anteriormente.

Ejemplo 1.1. Consideremos el siguiente modelo relacional. En el mismo se consideran cuatro mundos posibles (w_1, w_2, w_3 y w_4), los cuales cumplen con determinadas variables proposicionales (etiquetadas en los nodos) y pueden ser accesibles desde otros mundos. Por ejemplo, desde el mundo w_2 se puede acceder a w_3, mientras que a w_2 se lo puede acceder desde w_1. Por otro lado, w_3 puede accederse a sí mismo y a w_4.



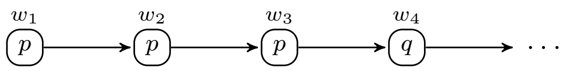
En este caso, la evaluación de las fórmulas se realiza desde un mundo en particular. Es decir, localmente. Si esta fórmula es meramente proposicional, la verificación se reduce a comprobar qué variables proposicionales se cumplen o no en el mundo actual. Por ejemplo, la fórmula $p \vee q$ es verdadera en los mundos w_1 y w_3, mientras que no lo es en w_2 y w_4. Sin embargo, al tener una fórmula con modalidades, el proceso requiere analizar los mundos que son accesibles desde el punto de evaluación, es decir, el conjunto de mundos considerados posibles. Luego procederemos a aplicar la interpretación de los operadoras modales correspondientes. Por ejemplo, si tomamos w_1 como punto de evaluación, podemos ver que $\Box r$, que se lee como “es necesario que r”, es verdadera dado que todos los mundos accesibles (o posibles) desde éste (w_2 y w_4) cumplen r.

Por otro lado, si tomamos w_3 como punto de evaluación, $\Box r$ no se cumple puesto que w_3 es un mundo accesible que no cumple r. Sin embargo, podemos ver que $\Diamond r$, que se lee como "es posible que r", es verdadera. Esto se debe a que existe un mundo w_4 accesible desde w_3 (un mundo posible) tal que se cumple r. $\Box$

La lógica que presentamos es conocida como Lógica Modal Básica, ya que es de alguna manera la extensión más pequeña (o una de ellas) que puede realizarse sobre la lógica proposicional, agregando nuevo poder expresivo. Además, es un resultado conocido que esta lógica es un fragmento estricto de la lógica de primer orden.

Como mencionamos anteriormente, de la misma manera en que podemos hablar de posibilidad o necesidad, también podemos definir operadores que describan otros modos de verdad. Una de las interpretaciones más populares debido a sus aplicaciones dentro de las ciencias de la computación es la "temporal". Las llamadas Lógicas Temporales [92] son lógicas modales que nos permiten expresar sentencias como "mañana lloverá" o "eventualmente saldrá el sol". Las lógicas temporales son útiles por ejemplo para caracterizar y verificar las especificaciones que un sistema tiene o debe tener al ejecutarse en un periodo de tiempo. En algunos sistemas, dichas ejecuciones se las puede visualizar como una sucesión discreta (o traza) de estados. Una de sus variantes es la llamada lógica lineal temporal (LTL) [32, 92], que cuenta con modalidades que describen cuándo una propiedad ψ se cumple en el siguiente estado ($\bigcirc\psi$), siempre en el futuro ($\Box\psi$), eventualmente en el futuro ($\Diamond\psi$), o hasta que se cumpla otra fórmula φ ($\psi \mathrm{U} \varphi$). El siguiente ejemplo ilustra el uso de LTL.

Ejemplo 1.2. Consideremos la siguiente traza de estados. La misma se representa mediante un grafo ligeramente diferente al visto en el Ejemplo 1.1. Dado que se está modelando una sucesión discreta de eventos que representa el paso del tiempo, el número de mundos a considerar es infinito numerable y la relación de accesibilidad es serial. Es decir, todo estado tiene un sucesor vía dicha relación.



Evaluando desde el estado inicial w_1, se tiene que p se cumple en el mismo y en los siguientes ($\bigcirc p$ y $\bigcirc\bigcirc p$). Sin embargo, p no se cumple globalmente, dado que en w_3 se cumple únicamente q ($\neg\Box p$). Por otro lado, p se cumple hasta que q lo haga en el estado w_4 ($p \mathrm{U} q$). Finalmente, se puede decir que q eventualmente se cumple ($\Diamond q$). $\Box$

Si bien este tipo de modelos presentan cambios y restricciones comparado con el Ejemplo 1.1, el mismo no altera la esencia de la lógica modal. Esto se debe a que por un lado sus operadores establecen modos de verdad, en este caso los referidos al tiempo, y los modelos son una clase particular de los modelos de Kripke introducidos anteriomente. Además, los operadores $\Diamond$ y $\Box$ siguen teniendo la misma semántica a pesar de tener diferentes lecturas. En el plano temporal, existen también variantes que consideran otras nociones de tiempo. Por ejemplo, las lógicas de árboles de computación (CTL y su extensión CTL*), que consideran múltiples ramas de ejecución en vez de una interpretación lineal de las mismas [33].

Si bien las lógicas temporales nos permiten describir cierto comportamiento de un programa o pieza de software de manera compacta, también puede resultar necesario ser más específicos en cuanto a las instrucciones y al flujo de ejecución de las mismas, sin contar necesariamente con una dimensión temporal. La lógica proposicional dinámica (PDL) [26, 45, 55, 104], fue diseñada para razonar y describir programas. Para ello, cuenta con operadores de composición (;) de instrucciones, ejecución no determinística ($\cup$) de las mismas, repetición de un comando particular (*), y consulta acerca de si una propiedad se cumple o no en un estado (?). Con estos podemos abstraernos de un problema particular con el que estamos trabajando, o definir secuencias de acciones lo suficientemente cercanas y fieles a la sintaxis de un lenguaje determinado, como podemos visualizar en el siguiente ejemplo.

Ejemplo 1.3. Consideremos el siguiente pseudocódigo con dos variables, x e y, que admiten sólo dos valores posibles, 0 ó 1.

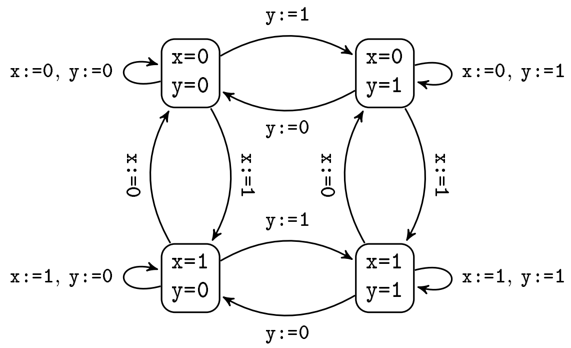
```
if (y == 0)
    then x:=1
else
    x:=0
```

Si el valor de `y` es 0, se le asigna a `x` el valor 1. En caso contrario, se le asigna 0. En PDL podemos representar la expresión `if-then-else` de la forma

$$\alpha = (((\texttt{y == 0})?; \texttt{x:=1}) \cup ((\neg(\texttt{y == 0}))?; \texttt{x:=1})).$$

Es decir, en un determinado estado, al ejecutar α, se ejecuta de manera no determinística $((\texttt{y == 0})?; \texttt{x:=1})$ ó $((\neg(\texttt{y == 0}))?; \texttt{x:=1})$, mediante la operación $\cup$ entre ambas. Como la variable `y` cumplirá únicamente una de las dos condiciones (es igual a 0 ó no lo es), sólo uno de los dos programas será ejecutado y realizará el cambio representado en el código de arriba. El test de igualdad a 0 se hace mediante la instrucción `y==0`?, que se compone con la asignación `x:=0` en caso de que corresponda. Para el caso distinto a 0, la sintaxis es análoga.

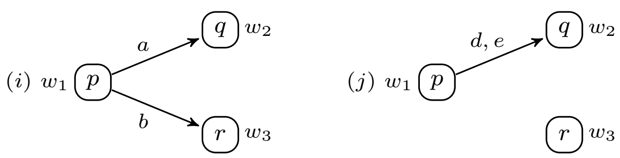
Nuevamente, el flujo de ejecución de α es definido mediante un grafo dirigido, siendo los nodos los estados que representan los valores de las variables y las aristas la ejecución de las asignaciones. En este caso, la relación de accesibilidad posee etiquetas con las diferentes operaciones que representa. Podemos ver por ejemplo que desde el estado donde $x = 0$ e $y = 0$ (estado de arriba a la izquierda), si le asignamos a y el valor 1, obtenemos el estado de arriba a la derecha, donde $y = 1$ y x mantiene su valor 0.



En este modelo, se puede verificar que, no importa desde cuál estado se empiece, al ejecutar el programa α, se terminan en estados en los que las variables `x` e `y` tiene valores diferentes. En PDL, esto se puede expresar que todos los estados cumplen que $[\alpha](x = 0 \leftrightarrow y = 1)$, que quiere decir que "después de ejecutar α desde un estado, x e y tienen valores diferentes". $\square$

En el Ejemplo 1.3 hablamos de relaciones de accesibilidad etiquetadas. Más específicamente sobre aristas que denotan diferentes tipos acciones (asignaciones) ejecutables por parte de un sistema. Sin embargo, en lógicas modales podemos tomar un paso más de abstracción sobre este sistema ejecutor y considerar sujetos capaces de ejecutar ciertas instrucciones. En general, esto se modela mediante el concepto de *agente*, una entidad autónoma abstracta capaz de efectuar diferentes acciones por cuenta propia (ya sea un ser humano, un robot, etcétera), teniendo en cuenta muchas veces su percepción sobre el entorno. Más aún, se pueden considerar múltiples agentes y modelar interacciones entre estos como, por ejemplo, acciones colectivas. Tal es el caso de la lógica de coaliciones (CL) que razona sobre la cooperación, las estrategias y las acciones por parte de agentes que se organizan en grupos o coaliciones [65, 88]. Dada una coalición de agentes $C \subseteq \mathsf{Agt}$, se define un operador modal $[C]$ para representar que existe una acción colectiva para C tal que una vez ejecutada, cumplen con una determinada propiedad, independientemente de la acción que tomen los agentes que estén por fuera de C.

Ejemplo 1.4. Tengamos en cuenta un conjunto de agentes $\mathsf{Agt} = \{i, j\}$ que deben tomar una decisión en conjunto. Dependiendo del estado, cada agente puede ejecutar una determinada acción. Esto se puede representar mediante un grafo para cada agente de Agt, los cuales tienen los mismos estados pero diferentes aristas etiquetadas según las acciones que tienen a su disposición. Consideremos los siguientes grafos para i y j, respectivamente.

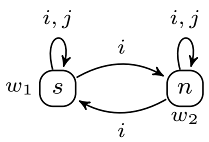


Podemos ver que, desde el estado w_1, todos los agentes pueden decidir moverse colectivamente a un estado w_2 donde se cumple q, con una acción a para el agente i, y con cualquiera de las acciones d ó e para el agente j ($[\mathsf{Agt}]q$). Sin embargo, como el agente j no tiene ninguna acción para moverse a w_3, los agentes no pueden tomar una acción colectiva desde w_1 que pueda llegar a un estado donde se cumpla r ($\neg[\mathsf{Agt}]r$). Más aún, esto nos dice que no importa qué acción realice i para llegar a un mundo donde se cumpla r, siempre existirá una acción por parte del agente j tal que llegará a un mundo donde no se cumpla ($\neg[\{i\}]r$). $\square$

Además de las características recién expuestas, CL puede extenderse en este caso con una dimensión temporal, tomando operadores similares a los de las lógicas temporales. Algunas de estas extensiones se denominan ATL y ATL* [3].

Con lo que hemos visto hasta ahora, el enfoque modal para representar este tipo de conceptos nos da un cierto grado de flexibilidad necesaria para ajustar las componentes de nuestra lógica. Considerando los agentes y las acciones, es posible también describir y razonar acerca de las habilidades de los mismos. Es decir, los cursos de acción que llevan a uno o varios agentes a lograr un objetivo. Este concepto será de gran interés en este trabajo, ya que estudiaremos este tipo de característica. Sin embargo, el otro aspecto importante será el estudio de la conciencia o conocimiento que un agente tiene acerca de sus propias capacidades. Por ejemplo, está claro que un agente es capaz de seguir un mapa y llegar a un punto de referencia en el mismo. Sin embargo, antes de tener conocimiento del mapa, el agente no *sabe* de la existencia de dicho camino, o más aún, como veremos más adelante, *no sabe cómo* llegar a destino a menos que tome conciencia del camino. El *conocimiento* de un agente usualmente puede representarse mediante las Lógicas Modales Epistémicas [62]. En particular, la lógica epistémica clásica es una lógica modal orientada a especificar el conocimiento que tienen los agentes sobre determinados hechos proposicionales. Por el lado sintáctico, para modelar este tipo de situaciones se indexan las modalidades con símbolos representando a cada agente. Sea $i \in \mathsf{Agt}$, se define el operador modal K_i, tal que $\mathsf{K}_i\psi$ se lee como "el agente i sabe que ψ se cumple". Mientras tanto, por el lado semántico, la representación de este conocimiento se basa en una relación de indistinguibilidad para cada agente. Dicha indistinguibilidad permite describir, para cada agente, qué mundos considera como igualmente posibles desde su perspectiva. El operador de conocimiento (K_i) en este caso se comporta como el de necesidad ($\Box$) de la lógica modal básica. Un agente determinado sabe un hecho proposicional desde un mundo en particular si todos los mundos indistinguibles de este cumplen dicha propiedad.

Ejemplo 1.5. Consideremos dos agentes i y j tales que se encuentran dentro de una casa. El agente i se encuentra lejos de cualquier ventana o abertura. Por ende, desde su perspectiva, no distingue si el día está soleado o nublado. En cambio, el agente j se encuentra al lado de la ventana. Por lo tanto, en cualquiera de las dos situaciones, si está nublado o soleado, el agente podrá distinguir ambos mundos entre sí. Un modelo que representa esta situación es el siguiente. Notar que las relaciones de indistinguibilidad están etiquetadas según el agente, y determinan su percepción del mundo.



Sean s y n las variables proposicionales que representan las situaciones *el día está soleado* y *el día está nublado* respectivamente. Dada la indistinguibilidad del agente i entre ambos escenarios (los considera igualmente posibles), desde el mundo w_1, i no sabe si el día está soleado o nublado ($\neg\mathsf{K}_i s$ y $\neg\mathsf{K}_i n$). Por otro lado, dado que el agente j es capaz de distinguirlos, el mismo sabe que está soleado ($\mathsf{K}_i s$), ya que s es verdadera en todos los mundos que considera posibles (en este caso, solamente w_1). $\square$

Este patrón de conocimiento se lo conoce como el "saber que", ya que considera el conocimiento del agente acerca de hechos proposicionales. Sin embargo, este patrón no es suficiente para describir todos los tipos de conocimiento que un agente puede o no tener. Por ejemplo, el "saber por qué" ciertos hechos ocurren, teniendo en cuenta las justificaciones para tal hecho. Otro ejemplo paradigmático, y de gran interés, es el patrón de conocimiento denominado del "saber cómo" [41]. Este patrón esta vinculado directamente a las habilidades del agente para lograr un objetivo. El interés por este concepto surge a partir de diferentes escenarios de las ciencias de la computación y de la inteligencia artificial, tales como los problemas de *planning* [49, 72].

Anteriormente consideramos lógicas que modelan el comportamiento de los agentes en lo referido a lograr un objetivo, y las acciones o habilidades que lo llevan a ello. En el último ejemplo consideramos la dimensión epistémica, en la cual se introduce el conocimiento de los agentes. Resulta natural entonces combinar ambos enfoques para modelar el conocimiento de los agentes acerca de sus propias habilidades. Sin embargo, los enfoques que consideran esta combinación resultan poco precisos a la hora de modelar específicamente el "saber cómo" de los agentes [59,66]. Es por ello que es necesario tener un enfoque más moderno para razonar sobre este patrón de conocimiento de las diferentes alternativas existentes.

En este trabajo nos enfocaremos en una lógica modal de saber cómo que represente este conocimiento mediante una relación de indistinguibilidad, esta vez no entre mundos, sino entre habilidades. En este sentido, en vez de considerar el conocimiento y las habilidades como dos dimensiones independientes, tomaremos las ideas de la lógica epistémica clásica para representar la incertidumbre de los agentes acerca de sus propias habilidades. Este enfoque resulta novedoso, no solo por los aspectos técnicos si no porque la lógica resultante, desde nuestra perspectiva, captura de manera más adecuada los principios subyacentes a este tipo de conocimiento. Además, será posible estudiar otras nociones epistémicas que no estaban presentes, como conocimiento colectivo, acciones que produzcan cambios de información en los agentes, y otras aplicaciones de este tipo de operaciones. Para ello, será necesario introducir primero algunos conceptos técnicos de lógica modal y de lógica epistémica, como así también uno de los enfoques más importantes para modelar el saber cómo [116,117]. El mismo será tomado como punto de partida para nuestro trabajo, por lo que luego discutiremos las principales diferencias y ventajas de la nueva lógica.

1.2. Una introducción formal a la lógica modal

En esta sección nos dedicaremos a introducir las definiciones formales que involucran a la lógica modal básica [24, 26], uno de los lenguajes modales más simples. En este sentido, introduciremos primero la sintaxis y semántica de la lógica. Luego, discutiremos algunas nociones importantes a la hora de estudiar y comprender una lógica, y que estarán presentes a lo largo de esta tesis. Entre ellas consideramos axiomatizaciones, bisimulaciones y la complejidad de tareas de inferencia.

Como se ha mencionado anteriormente, la sintaxis tiene como base la lógica proposicional (las variables proposicionales $p_1, p_2, \ldots$ con los conectores booleanos $\wedge$, $\vee$, $\neg$, $\rightarrow$ y $\leftrightarrow$), más dos modalidades $\Box$ y $\Diamond$. Estas representan los modos de verdad "es necesario que" y "es posible que" respectivamente. Por ejemplo, $\Box p_1$ y $\Diamond p_4$ se leen como "es necesario que p_1 se cumpla" y "es posible que p_4 se cumpla". Con esto, se pueden escribir fórmulas todavía más elaboradas como $\Box p_1 \rightarrow \Diamond p_1$ ("si es necesario que p_1 se cumpla, entonces es posible que p_1 se cumpla"), $\Box(p_2 \wedge p_3) \rightarrow \Box p_2$ ("si es necesario que p_1 y p_2 se cumplan, entonces es necesario que p_1 se cumpla") ó $\Diamond p_4 \rightarrow \Box\Diamond p_3$ ("si p_4 es posible, entonces es necesario que p_4 sea posible"). Formalmente, esta sintaxis se define de la siguiente manera.

Definición 1.1. Sea $\mathsf{Prop} = \{p_1, p_2, \ldots\}$ un conjunto numerable de símbolos de proposición. Las fórmulas de la lógica modal básica (ML) son las dadas por la gramática

$$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid \Box\varphi,$$

con $p \in \mathsf{Prop}$. Las constantes booleanas y demás operadores se definen como:

- $\top := (p \vee \neg p)$,
- $\bot := \neg\top$,
- $(\varphi \rightarrow \psi) := (\neg\varphi \vee \psi)$,
- $(\varphi \wedge \psi) := \neg(\neg\varphi \vee \neg\psi)$,
- $(\varphi \leftrightarrow \psi) := (\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi)$,
- $\Diamond\varphi := \neg\Box\neg\varphi$.

Notar que la definición de $\Diamond$ se da en términos de $\Box$. En el contexto de la lógica modal básica esto tiene sentido pues la expresión "es posible que p se cumpla" es equivalente a decir que "no es necesario que no se cumpla p". En estos casos, se tiene que $\Diamond$ es el "dual" de $\Box$ y viceversa.

Para el caso de la semántica, la misma se compone de dos partes: ***(1)*** la definición de los modelos a tratar y ***(2)*** la interacción de las fórmulas con dichos modelos. Esto último se da mediante la definición de una relación de satisfacción entre modelos y fórmulas, denotada como $\models$. Para la lógica modal básica, la semántica se introduce en términos de los llamados modelos de Kripke, estructuras relacionales compuestas por un conjunto no vacío de mundos, una relación de accesibilidad binaria entre estos y una función que describe qué variables proposicionales son verdaderas en cada mundo, es decir, son grafos dirigidos etiquetados. Formalmente:

Definición 1.2. Un modelo de Kripke definido sobre Prop es una tupla $\mathcal{M} = \langle \mathrm{W}, \mathrm{R}, \mathrm{V} \rangle$ tal que:

- W es un conjunto no vacío de mundos o estados (también denotado como $\mathrm{D}_{\mathcal{M}}$),
- $\mathrm{R} \subseteq \mathrm{W} \times \mathrm{W}$ es una relación de accesibilidad entre mundos, y
- $\mathrm{V} : \mathrm{W} \to 2^{\mathsf{Prop}}$ es una función de valuación que a cada mundo le asigna el conjunto de proposiciones que se cumplen en dicho mundo.

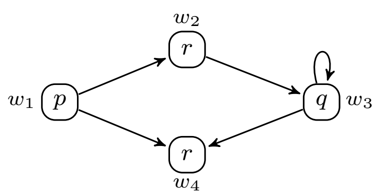
Sea $w \in \mathrm{W}$ decimos que $(\mathcal{M}, w)$ es un modelo punteado, y en general los paréntesis son omitidos.

Si bien la definición de V en este caso difiere de las introducidas en [24, 26] ($\mathrm{V}' : \mathsf{Prop} \to 2^{\mathrm{W}}$) y en [48] ($v : \mathsf{Prop} \times \mathrm{W} \to \{T, F\}$), claramente estas tres alternativas son equivalentes entre sí. Por otro lado, la introducción de los modelos punteados (en inglés, ***pointed models***) se debe a que las fórmulas de la lógica modal se interpretan localmente, es decir, con respecto a un punto de evaluación en el modelo. De esta manera la misma nos provee una perspectiva interna del mismo. Por último, estos modelos se representan gráficamente como un grafo dirigido en el cual los nodos son los mundos de W, cada uno etiquetado con las variables proposicionales que se cumplen vía V, y los arcos son los elementos de R.

Ejemplo 1.6. Se tiene $\mathcal{M} = \langle \mathrm{W}, \mathrm{R}, \mathrm{V} \rangle$ un modelo de Kripke tal que

- $\mathrm{W} = \{w_1, w_2, w_3, w_4\}$,
- $\mathrm{R} = \{(w_1, w_2), (w_1, w_4), (w_2, w_3), (w_3, w_4), (w_3, w_3)\}$, y
- $\mathrm{V}(w_1) = \{p\}$, $\mathrm{V}(w_2) = \mathrm{V}(w_4) = \{r\}$, $\mathrm{V}(w_4) = \{q\}$.

Su representación gráfica, vista en el Ejemplo 1.1, es la siguiente:



□

Finalmente, introducimos a continuación la relación de satisfacción $\models$.

Definición 1.3. Sean $\mathcal{M} = \langle \mathrm{W}, \mathrm{R}, \mathrm{V} \rangle$ un modelo de Kripke definido sobre Prop, y $w \in \mathrm{W}$. La relación de satisfacción $\models$ entre $\mathcal{M}$, w y las fórmulas de la lógica modal básica se define inductivamente de la siguiente manera:

$$
\begin{array}{lll}
\mathcal{M}, w \models p & \text{sii} & p \in \mathrm{V}(w), \\
\mathcal{M}, w \models \neg\varphi & \text{sii} & \mathcal{M}, w \not\models \varphi, \\
\mathcal{M}, w \models \varphi \vee \psi & \text{sii} & \mathcal{M}, w \models \varphi \text{ ó } \mathcal{M}, w \models \psi, \\
\mathcal{M}, w \models \Box\varphi & \text{sii} & \text{para todo } (w, v) \in \mathrm{R}, \text{ se tiene que } \mathcal{M}, v \models \varphi.
\end{array}
$$

Una fórmula φ es satisfacible sii $\mathcal{M}, w \models \varphi$ para algún $\mathcal{M} = \langle \mathrm{W}, \mathrm{R}, \mathrm{V} \rangle$ y $w \in \mathrm{W}$, y es válida (denotado como $\models \varphi$) sii $\mathcal{M}, w \models \varphi$ para todo $\mathcal{M} = \langle \mathrm{W}, \mathrm{R}, \mathrm{V} \rangle$ y $w \in \mathrm{W}$. Para un conjunto de fórmulas Ψ, $\mathcal{M}, w \models \Psi$ sii $\mathcal{M}, w \models \psi$ para todo $\psi \in \Psi$; y $\Psi \models \varphi$ sii para todo $\mathcal{M} = \langle \mathrm{W}, \mathrm{R}, \mathrm{V} \rangle$, $w \in \mathrm{W}$, $\mathcal{M}, w \models \Psi$ implica $\mathcal{M}, w \models \varphi$.

Ejemplo 1.7. Tomando el Ejemplo 1.6, $\mathcal{M}, w_1 \models \Box r$, pues para todo $(w_1, v) \in \mathrm{R}$, se tiene que $\mathcal{M}, v \models r$ con $v \in \{w_2, w_3\}$. Así, hay una correspondencia entre la modalidad “es necesario que r se cumpla” en el mundo w_1 con la idea que “todos los mundos accesibles desde w_1 cumplen r”. □

Axiomas:	TAUT	$\vdash \varphi$ para φ una tautología proposicional
	DIST	$\vdash \Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$
Reglas:	$\dfrac{\varphi \quad (\varphi \to \psi)}{\psi}$ MP	$\dfrac{\varphi}{\Box\varphi}$ NEC

Tabla 1.1: Sistema axiomático de una lógica modal normal **K**

La noción de fórmula válida (también llamada tautología) nos permite estudiar ciertas características de una lógica. Como es sabido, es posible caracterizar el comportamiento de una lógica mediante el conjunto de sus fórmulas válidas, es decir, de manera semántica. Sin embargo, en general resulta adecuado contar con una caracterización sintáctica de la lógica, mediante la noción de derivación en un sistema de prueba. En la Tabla 1.1, introducimos un sistema axiomático "a la Hilbert" para la lógica modal básica [48], denominado **K**. La misma se compone de todos los axiomas de la lógica proposicional (TAUT) y el axioma de distributividad de $\Box$ con respecto a la implicación (DIST). Además, se introducen dos reglas: una de Modus Ponens (MP) y otra de necesitación (NEC), en la cual si φ es un teorema, entonces $\Box\varphi$ también es un teorema. Toda lógica modal que tenga estos axiomas y reglas para cada operador se la considera como una lógica modal normal. La relación de derivación sintáctica entre conjuntos de fórmulas y fórmulas ($\Gamma \vdash \varphi$) se define de la manera usual.

Dado el sistema axiomático **K**, existen dos propiedades que vinculan las derivaciones sintácticas que pueden realizarse en **K** con las fórmulas que son semánticamente inferibles en los modelos. La primera es *completitud fuerte*, que establece que toda fórmula φ que sea consecuencia lógica de un conjunto de fórmulas Γ ($\Gamma \models \varphi$) es deducible del mismo conjunto ($\Gamma \vdash \varphi$). Mientras tanto, la segunda se llama *correctitud (o corrección)* y es la conversa: toda fórmula φ que sea deducible de un conjunto de fórmulas Γ ($\Gamma \vdash \varphi$) es consecuencia lógica del mismo conjunto ($\Gamma \models \varphi$). Con estas definiciones, se sabe que **K** cumple con estas propiedades.

Teorema 1.1 ([24,26])**.** *El sistema **K** es correcto y fuertemente completo con respecto a la clase de todos los modelos de Kripke.*

La presentación del sistema axiomático para la Lógica Modal Básica merece algunos comentarios. Primero que nada, vale la pena destacar que la misma se realiza mediante los llamados *axiomas esquema*. Con esto queremos decir que cada uno de los axiomas y reglas no son solo válidos para una fórmula en particular. En dicha presentación, las expresiones φ y ψ representan variables de fórmulas, por lo que los axiomas y reglas valen para cualquier expresión que cumpla con los patrones descriptos en dichos axiomas esquema. Por ejemplo, por el axioma DIST, se tiene que $\Box(p \to q) \to (\Box p \to \Box q)$ es un teorema, y que también lo es $\Box((p \wedge q) \to r) \to (\Box(p \wedge q) \to \Box r)$. Una presentación alternativa de este tipo de sistemas, consiste en introducir axiomas y reglas únicas con respecto a un símbolo proposicional, y luego una regla de *sustitución uniforme*. Por ejemplo, el axioma DIST puede fijarse a $\Box(p \to q) \to (\Box p \to \Box q)$, para $p, q \in \mathsf{Prop}$ fijos, con una regla de sustitución que establece que si φ es un teorema, luego cualquier reemplazo sintáctico simultáneo de alguna de sus variables proposicionales (p) por una fórmula arbitraria (ψ) también es un teorema (denotado $\varphi[p := \psi]$). Los axiomas esquema y la sustitución uniforme son mecanismos fundamentales para generar nuevos teoremas, y resultan equivalentes para el caso del sistema **K**. Sin embargo, como veremos más adelante, esta situación no siempre sucede. Particularmente, un caso que resulta de nuestro interés es el de las lógicas modales que incluyen operadores dinámicos, que modifican el modelo a medida que se evalúa una fórmula, donde la propiedad de sustitución falla [63]. Una estrategia para lidiar con estas complicaciones suele ser mediante los llamados *axiomas de reducción* [106], un conjunto de equivalencias entre fórmulas, cuya aplicación sucesiva permite la eliminación de los operadores dinámicos. Una vez eliminados, se puede utilizar el sistema base de la lógica sin estos operadores para obtener un sistema completo. Esto es posible en casos donde la expresividad del operador dinámico en cuestión puede ser captada por los demás operadores del lenguaje base. Si bien la propiedad de sustitución falla en muchos casos, los axiomas de reducción nos permiten eliminar los operadores que causan problemas para tal propiedad, y trabajar con un fragmento donde la sustitución o los axiomas esquema funcionen. Por este motivo resulta de interés determinar si propiedades como la sustitución falla, y las maneras de subsanar este hecho. En casos donde el lenguaje base no permite caracterizar los operadores dinámicos (es decir, donde estos últimos son más expresivos) es posible también agregar expresividad al lenguaje base con operadores que se comporten de la manera esperada (por ejemplo, satisfaciendo sustitución) y de esta manera obtener la expresividad necesaria para caracterizar los operadores dinámicos. Daremos detalles concretos de este tipo de argumentos más adelante en esta tesis.

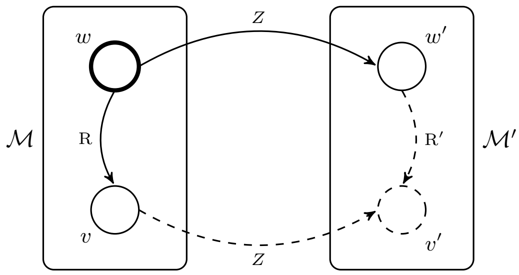
Otro aspecto importante a la hora de comprender las lógicas modales es el concepto de bisimulaciones, que busca comparar y analizar el comportamiento entre dos modelos distintos, y capturar la expresividad del lenguaje subyacente [95,96]. En síntesis, dos modelos son bisimilares si la estructura de ambos comparte ciertas características desde el punto de vista de la lógica. Una bisimulación relaciona mundos entre dos modelos, tal que comparten los mismos símbolos de proposición, y que además si en uno de ellos se puede atravesar un eje, en el otro también es posible y a su vez los mundos de destino por este eje deben ser bisimilares. A continuación introducimos formalmente esta noción.

Definición 1.4 (Bisimulación)**.** Sean $\mathcal{M} = \langle \mathrm{W}, \mathrm{R}, \mathrm{V} \rangle$ y $\mathcal{M}' = \langle \mathrm{W}', \mathrm{R}', \mathrm{V}' \rangle$ dos modelos de Kripke. Una relación no vacía $Z \subseteq \mathrm{W} \times \mathrm{W}'$ es una bisimulación entre $\mathcal{M}$ y $\mathcal{M}'$ si y sólo si para todo wZw' implica lo siguiente.

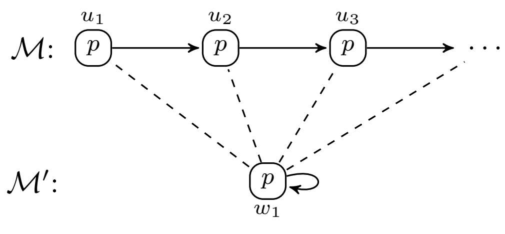
- **Atom**: $\mathrm{V}(w) = \mathrm{V}'(w')$.
- **Zig**: si $(w, v) \in \mathrm{R}$, entonces existe $v' \in \mathrm{W}'$ tal que vZv' y $(w', v') \in \mathrm{R}'$.
- **Zag**: si $(w', v') \in \mathrm{R}'$, entonces existe $v \in \mathrm{W}$ tal que vZv y $(w, v) \in \mathrm{R}$.

Se escribe $\mathcal{M}, w \leftrightarroweq \mathcal{M}', w'$ cuando hay una bisimulación Z entre $\mathcal{M}$ y $\mathcal{M}'$ tal que wZw'.

El siguiente diagrama muestra las condiciones que **Zig** describe. Sea $w \in \mathrm{W}$, supongamos que wZw' y $(w, v) \in \mathrm{R}$. Luego, por **Zig**, necesariamente existe un $v' \in \mathrm{W}$ tal que $(w', v') \in \mathrm{R}'$ y vZv', completando así el diagrama. Notar que el caso de **Zag** es simétrico, basta con espejar la imagen en el eje vertical y se tiene el diagrama equivalente.



Ejemplo 1.8. Sean $\mathcal{M} = \langle \mathrm{W}, \mathrm{R}, \mathrm{V} \rangle$ y $\mathcal{M}' = \langle \mathrm{W}', \mathrm{R}', \mathrm{V}' \rangle$ los siguientes modelos.



Tomando $Z = \{(u_1, w_1), (u_2, w_1), (u_3, w_1), (u_4, w_1), \ldots\}$, representado en el gráfico con las líneas punteadas, se puede observar que Z cumple **Atom**, dado que todos los mundos cumplen las mismas variables proposicionales. Para probar **Zig**, sean $u_i Z w_1$ y $(u_i, u_{i+1}) \in \mathrm{R}$, existe $v' = w_1$ tal que cumple las condiciones de la definición. Con un razonamiento similar se obtiene **Zag**. Con esto, Z es una bisimulación entre $\mathcal{M}$ y $\mathcal{M}'$. Además, $\mathcal{M}, u_i \leftrightarroweq \mathcal{M}', w_1$ para todo $i = 1, 2, \ldots$, dado que $u_i Z w_1$. De este ejemplo, se puede concluir que, en algunos casos, la lógica modal básica no distingue un modelo de un solo elemento de uno infinito. $\square$

La definición de bisimilaridad está conectada con la de equivalencia de fórmulas. Esto es debido a que, dados dos modelos que se comportan de manera similar desde la perspectiva de una lógica modal determinada, deben cumplir las mismas propiedades desde esa lógica. Estas propiedades nos permiten comprender el poder expresivo de la lógica.

Definición 1.5 (Equivalencia)**.** Sean $\mathcal{M} = \langle \mathrm{W}, \mathrm{R}, \mathrm{V} \rangle$ y $\mathcal{M}' = \langle \mathrm{W}', \mathrm{R}', \mathrm{V}' \rangle$ dos modelos, $w \in \mathrm{W}$ y $w' \in \mathrm{W}$ mundos. $\mathcal{M}, w$ y $\mathcal{M}', w'$ son *equivalentes* ($\mathcal{M}, w \leftrightsquigarrow \mathcal{M}', w'$) si y sólo si, para cada fórmula φ,

$$\mathcal{M}, w \models \varphi \quad \text{sii} \quad \mathcal{M}', w' \models \varphi.$$

Con esto se puede establecer la correspondencia entre $\underline{\leftrightarrow}$ y $\leftrightsquigarrow$.

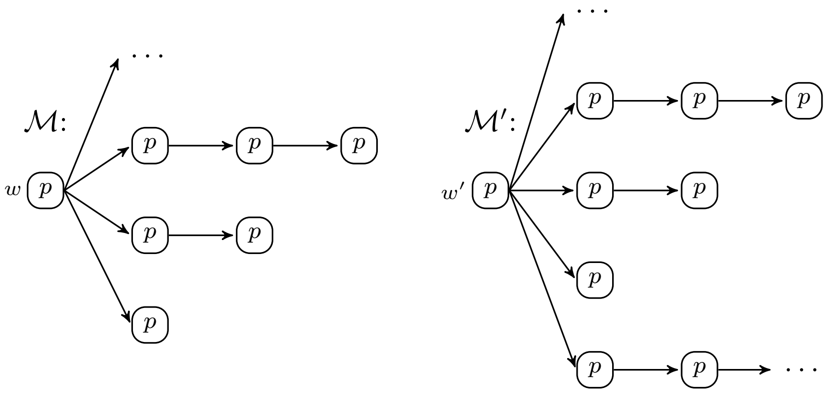
Teorema 1.2 (Invarianza bajo bisimulaciones [26])**.** *Sean* $\mathcal{M} = \langle \mathrm{W}, \mathrm{R}, \mathrm{V} \rangle$ *y* $\mathcal{M}' = \langle \mathrm{W}', \mathrm{R}', \mathrm{V}' \rangle$ *dos modelos,* $w \in \mathrm{W}$ *y* $w' \in \mathrm{W}'$ *mundos. Entonces,*

$$\mathcal{M}, w \underline{\leftrightarrow} \mathcal{M}', w' \quad \textit{implica} \quad \mathcal{M}, w \leftrightsquigarrow \mathcal{M}', w'.$$

Ejemplo 1.9. Tomando los modelos del Ejemplo 1.8, como $\mathcal{M}, u_i \underline{\leftrightarrow} \mathcal{M}', w_1$, entonces $\mathcal{M}, u_i \leftrightsquigarrow \mathcal{M}', w_1$. □

Desafortunadamente, la vuelta no siempre se cumple para modelos arbitrarios.

Ejemplo 1.10 ([26])**.** Consideremos dos modelos $\mathcal{M} = \langle \mathrm{W}, \mathrm{R}, \mathrm{V} \rangle$ y $\mathcal{M}' = \langle \mathrm{W}', \mathrm{R}', \mathrm{V}' \rangle$, $p \in \mathsf{Prop}$ tal que ambos modelos tienen infinitas ramas desde sus respectivos nodos raíz. Supongamos además que todas las ramas de $\mathcal{M}$ son de largo finito mientras que $\mathcal{M}'$ tiene además una rama de largo infinito.



Se puede ver que $\mathcal{M}, w \leftrightsquigarrow \mathcal{M}', w'$. Sin embargo, la rama infinita de $\mathcal{M}'$ imposibilita una bisimulación que vincule a sus nodos raíces, dado que la misma no podría emparejarse con ninguna de las ramas de largo finito en $\mathcal{M}$. Por lo tanto, $\mathcal{M}, w \underline{\leftrightarrow} \mathcal{M}', w'$ no sucede. □

Para que se cumpla la conversa, usualmente es necesario restringir la clase de modelos con los cuales se trabaja. Las clases de modelos que satisfacen la vuelta del Teorema 1.2 se conocen como clases de Hennessy-Milner. En el caso de las lógicas modales, normalmente esta clase es la de los modelos con imagen finita [26]. Es decir, modelos tales que por cada mundo, se tiene una cantidad finita de sucesores.

Teorema 1.3 (Equivalencia implica bisimilaridad [26])**.** *Sean* $\mathcal{M} = \langle \mathrm{W}, \mathrm{R}, \mathrm{V} \rangle$ *y* $\mathcal{M}' = \langle \mathrm{W}', \mathrm{R}', \mathrm{V}' \rangle$ *modelos con imagen finita,* $w \in \mathrm{W}$ *y* $w' \in \mathrm{W}'$ *mundos. Entonces,*

$$\mathcal{M}, w \leftrightsquigarrow \mathcal{M}', w' \quad \textit{implica} \quad \mathcal{M}, w \underline{\leftrightarrow} \mathcal{M}', w'.$$

De esta forma, las bisimulaciones son una herramienta relevante para entender el poder expresivo de una lógica modal. Más aún, el resultado de estos teoremas facilita la tarea de comparar la expresividad entre dos lógicas modales, especialmente cuando se analizan extensiones de un lenguaje base con nuevos operadores. Esto es relevante dado que, en algunos casos queremos saber qué operadores de una lógica pueden ser expresados en términos de otros o no. Para tener una definición formal, tomaremos la notación que provee [11, Definición 2.8].

Definición 1.6 ($\mathsf{L}_1 \preceq \mathsf{L}_2$)**.** Sean L_1 y L_2 dos lógicas, L_1 es tan expresiva como L_2 o reducible a L_2 (denotado como $\mathsf{L}_1 \preceq \mathsf{L}_2$) si para toda fórmula φ de L_1 existe una fórmula φ' de L_2 tal que para todo modelo $\mathcal{M}$ y $w \in \mathrm{D}_{\mathcal{M}}$, se cumple que

$$\mathcal{M}, w \models \varphi \;\; \text{sii} \;\; \mathcal{M}, w \models \varphi'.$$

Caso contrario, $\mathsf{L}_1 \not\preceq \mathsf{L}_2$. Si $\mathsf{L}_1 \preceq \mathsf{L}_2$ pero $\mathsf{L}_2 \not\preceq \mathsf{L}_1$, se dice que L_2 es más expresiva que L_1 y se escribe $\mathsf{L}_1 \prec \mathsf{L}_2$. Si $\mathsf{L}_1 \preceq \mathsf{L}_2$ y $\mathsf{L}_2 \preceq \mathsf{L}_1$, entonces son equivalentes y se lo denota como $\mathsf{L}_1 \equiv \mathsf{L}_2$.

Para garantizar que $\mathsf{L}_1 \preceq \mathsf{L}_2$, el procedimiento usual es dar una función de traducción $T : \mathsf{L}_1 \to \mathsf{L}_2$ tal que para cada fórmula φ de L_1 se le asigna su fórmula equivalente $T(\varphi)$ de L_2 en la clase de modelos considerada. Con esto, se busca que para todo modelo $\mathcal{M}$ y $w \in \mathrm{D}_{\mathcal{M}}$,

$$\mathcal{M}, w \models \varphi \ \text{ sii } \ \mathcal{M}, w \models T(\varphi).$$

Ejemplo 1.11. Consideremos la lógica $\mathsf{L}_{\triangle}$ definida por la siguiente gramática

$$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid \triangle\varphi,$$

y que además la semántica del operador $\triangle$ sea

$$\mathcal{M}, w \models \triangle\varphi \quad \text{sii} \quad \text{no existe } (w, v) \in \mathrm{R} \text{ tal que } \mathcal{M}, v \models \varphi.$$

Claramente, $\triangle\varphi$ es semánticamente equivalente a $\Box\neg\varphi$. Por lo tanto, basta con definir inductivamente la función de traducción $T : \mathsf{L}_{\triangle} \to \mathsf{ML}$ tal que

- $T(p) = p$,
- $T(\neg\varphi) = \neg T(\varphi)$,
- $T(\varphi \vee \varphi) = T(\varphi) \vee T(\varphi)$, y
- $T(\triangle\varphi) = \Box\neg T(\varphi)$.

Luego, $\mathsf{L}_{\triangle} \preceq \mathsf{ML}$, dado que la traducción vale para todo modelo de Kripke. Más aún, se tiene que $\mathsf{ML} \preceq \mathsf{L}_{\triangle}$ utilizando una traducción $T' : \mathsf{ML} \to \mathsf{L}_{\triangle}$ tal que $T'(\Box\varphi) = \triangle\neg\varphi$. Con esto, $\mathsf{L}_{\triangle} \equiv \mathsf{ML}$. $\Box$

Para demostrar lo contrario, es decir que $\mathsf{L}_1 \not\preceq \mathsf{L}_2$, el procedimiento usual es tomar dos modelos $\mathcal{M}$ y $\mathcal{M}'$ que cumplan las mismas fórmulas de L_2 (usualmente bisimilares en L_2 y luego por Teorema de Invarianza sabemos que satisfacen las mismas fórmulas), y dar una fórmula ψ de L_1 tal que se cumpla en uno de los modelos ($\mathcal{M}, w \models \psi$) mientras que en el otro no ($\mathcal{M}', w \not\models \psi$). Ello invalida la existencia de una función de traducción $T : \mathsf{L}_1 \to \mathsf{L}_2$ ya que de existir una, ψ sería traducible y se tendría la siguiente serie de equivalencias:

$$\mathcal{M}, w \models \psi \ \text{ sii } \ \mathcal{M}, w \models T(\psi) \ \text{ sii } \ \mathcal{M}', w \models T(\psi) \ \text{ sii } \ \mathcal{M}', w \models \psi.$$

Y con esto, $\mathcal{M}', w \models \psi$, contradiciendo la hipótesis sobre ψ.

Ejemplo 1.12. Consideremos una extensión de ML con una modalidad de diferencia, $\mathsf{ML} + \langle\neq\rangle$, definida por la siguiente gramática

$$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid \Box\varphi \mid \langle\neq\rangle\varphi,$$

extendiendo la semántica de $\models$ al operador $\langle\neq\rangle$ de la forma

$$\mathcal{M}, w \models \langle\neq\rangle\varphi \quad \text{sii} \quad \text{existe } v \neq w \text{ tal que } \mathcal{M}, v \models \varphi.$$

Dado que $\mathsf{ML} + \langle\neq\rangle$ es una extensión de ML, es fácil ver que $\mathsf{ML} \preceq \mathsf{ML} + \langle\neq\rangle$ utilizando una función de traducción tal que $T(\varphi) = \varphi$. Por otro lado, se puede demostrar que $\mathsf{ML} + \langle\neq\rangle \not\preceq \mathsf{ML}$ tomando los modelos del Ejemplo 1.8. Dado que son bisimilares ($\mathcal{M}, u_1 \leftrightarroweq \mathcal{M}', w_1$), se tiene que $\mathcal{M}, u_1 \leftrightsquigarrow \mathcal{M}', w_1$. Sin embargo, existe una fórmula, $\langle\neq\rangle p$, tal que $\mathcal{M}, u_1 \models \langle\neq\rangle p$ (pues $\mathcal{M}$ tiene más de un mundo cumpliendo p), pero que $\mathcal{M}', w_1 \not\models \langle\neq\rangle p$. Esto quiere decir que $\mathsf{ML} + \langle\neq\rangle$ puede distinguir modelos de más de un elemento, algo que ML no es capaz. Por lo tanto, $\mathsf{ML} + \langle\neq\rangle \not\preceq \mathsf{ML}$ y con esto $\mathsf{ML} \prec \mathsf{ML} + \langle\neq\rangle$. $\Box$

Como mencionamos anteriormente, las lógicas modales se caracterizan en general por sus buenas propiedades computacionales. En este sentido, usualmente es importante estudiar la complejidad computacional de dos problemas de decisión principales: el problema model-checking (o verificación de modelos) y el problema de satisfacibilidad. Dichos problemas tienen diversas aplicaciones en el modelado y verificación de software. El primero se formula como "dado un modelo y una fórmula de la lógica, determinar si dicha fórmula se cumple o no en el modelo" mientras que el segundo es de la forma "dada una fórmula de la lógica, determinar si existe un modelo que la satisfaga". La clase de complejidad computacional a la que pertenecen ambos problemas varía según la lógica que se esté estudiando. Por ejemplo, para la lógica proposicional, se tienen como estructuras las valuaciones de las variables proposicionales. Dada una fórmula, verificar que una valuación la satisface es polinomial con respecto al tamaño de la fórmula, mientras que para determinar si es satisfacible, es necesario tomar decisiones no determinísticas. Considerando esto, se tienen los siguientes resultados computacionales.

Teorema 1.4 ([12, 35]). *El problema de model-checking para la lógica proposicional está en* P. *El problema de satisfacibilidad es* NP*-completo.*

Para el caso de la lógica modal básica, al poder hablar sobre mundos posibles con diferentes valuaciones, esto impacta en los problemas de decisión. Para el caso de satisfacibilidad es posible forzar estructuras con un número exponencial de mundos. Sin embargo, es suficiente con contar un espacio de almacenamiento polinomial con respecto al tamaño de la fórmula.

Teorema 1.5 ([25]). *El problema de model-checking para la lógica modal básica está* P. *El problema de satisfacibilidad es* PSpace*-completo.*

Para ilustrar el buen comportamiento de la lógica modal básica introducimos a continuación los resultados correspondientes para la lógica de primer orden. Es un resultado conocido que la lógica modal básica es un fragmento de la lógica de primer orden.

Teorema 1.6 ([25]). *El problema de model-checking para la lógica de primer orden es* PSpace*-completo. El problema de satisfacibilidad es indecidible.*

Con estos resultados, se puede inferir una correlación: cuanto más expresiva sea una lógica, sus problemas de model-checking y satisfacibilidad pertenecerán a una clase de complejidad más alta (a menos que influyan otros factores como la posibilidad de escribir fórmulas de manera más concisa). Por ello se debe tener cuidado a la hora de considerar una lógica modal desde un punto de vista computacional.

Teniendo en cuenta todos estos aspectos, veremos más adelante que las lógicas modales resultan ser una buena alternativa para modelar sistemas multiagente. Una buena definición tanto sintáctica como semántica es capaz de proveer un lenguaje legible y que exprese las propiedades que se buscan, con una complejidad computacional deseable. Además, las lógicas modales pueden proveer herramientas uniformes que nos permitan comprender el comportamiento de los lenguajes en cuestión. Nos abocaremos a tomar esta perspectiva en los capítulos subsiguientes.

Capítulo 2

Lógicas epistémicas

> Saber que sabemos lo que sabemos, y saber que no sabemos lo que no sabemos, eso es verdadero conocimiento.
>
> Confucio

2.1. Lógicas epistémicas

Originada a principios de los años 1960s [62, 108], la lógica epistémica es un subcampo de la epistemología que se ocupa de los enfoques lógicos del conocimiento, la creencia y conceptos similares [93]. Particularmente, la lógica epistémica clásica es un formalismo lógico que permite razonar sobre el conocimiento de un agente o un conjunto de agentes mediante una noción de incertidumbre acerca de los hechos que estos perciben acerca del mundo real. La lógica epistémica ha contribuido en el estudio formal de nociones espistémicas en diveros campos, tales como la filosofía [58], las ciencias de la computación [39, 81] y la economía [90], entre otras.

Las lógicas epistémicas generalmente se definen en el contexto de las lógicas modales. Los enfoques más tradiciones suelen representar la noción de *saber qué*. Es decir, se enfocan en el estudio de sentencias de la forma *"el agente sabe que está soleado en París"* o *"el robot sabe que está al lado de una pared"*. Introduciremos a continuación los principales elementos de la lógica epistémica clásica. Como podremos observar, la misma es una lógica modal con la particularidad de que en vez de una sola modalidad $\Box$, contamos con múltiples modalidades K_i, siendo cada una indexada por un símbolo de agente i de un conjunto Agt previamente definido. Cada modalidad será interpretada sobre una relación de accesibilidad correspondiente en el modelo. En el resto de la tesis, definimos a Agt como un conjunto finito de agentes.

Definición 2.1. Las fórmulas de la lógica epistémica estándar L_K son las dadas por la siguiente gramática

$$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid \mathsf{K}_i\varphi,$$

con $p \in \mathsf{Prop}$ e $i \in \mathsf{Agt}$. Las constantes booleanas y demás operadores se definen como:

- $\top := (p \vee \neg p)$,
- $\bot := \neg\top$,
- $(\varphi \to \psi) := (\neg\varphi \vee \psi)$,
- $(\varphi \wedge \psi) := \neg(\neg\varphi \vee \neg\psi)$,
- $(\varphi \leftrightarrow \psi) := (\varphi \to \psi) \wedge (\psi \to \varphi)$,
- $\widehat{\mathsf{K}}_i\varphi := \neg\mathsf{K}_i\neg\varphi$ (dual de K_i).

Las fórmulas de la forma $\mathsf{K}_i\varphi$ se leen como *"el agente i sabe que φ es verdadera"*, mientras que $\widehat{\mathsf{K}}_i\varphi$ se lee como *"el agente i considera posible que φ sea verdadera"*.

A nivel semántico, las fórmulas epistémicas también se interpretan sobre modelos de Kripke [24, 26]. En estos modelos, el conocimento de cada agente está determinado por una relación epistémica de *indistinguibilidad* entre mundos, denotada como $\sim_i$, y representa la incertidumbre que el agente tiene entre diferentes situaciones. Los mundos que estén relacionados por $\sim_i$ para un determinado agente i, son considerados "indistinguibles" desde su perspectiva. Usualmente, para capturar las propiedades del conocimiento, se asume que $\sim_i$ es una relación de equivalencia [24, 26]. Esto es natural pues ***(1)*** el agente no distingue un mundo de sí mismo (reflexividad); ***(2)*** si no distingue un mundo de otro, entonces no distingue el otro del uno (simetría); ***(3)*** y si no distingue un primer mundo de un segundo y un segundo de un tercero, entonces no va a distinguir el primero del tercero (transitividad). Con esto, dado un conjunto de agentes Agt, los modelos de esta lógica son de la forma $\mathcal{M} = \langle \mathrm{W}, \{\sim_i\}_{i\in\mathsf{Agt}}, \mathrm{V}\rangle$, siendo cada $\sim_i$ una relación de equivalencia. Formalmente, se tiene la siguiente definición.

Definición 2.2. Un modelo de Kripke epistémico multiagente, definido sobre un conjunto de variables proposicionales Prop y un conjunto de agentes Agt, es una tupla $\mathcal{M} = \langle \mathrm{W}, \{\sim_i\}_{i\in\mathsf{Agt}}, \mathrm{V}\rangle$ tal que:

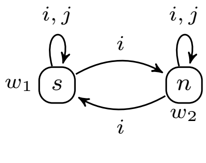
- W es un conjunto no vacío de mundos (también denotado como $\mathrm{D}_{\mathcal{M}}$),
- $\sim_i \subseteq \mathrm{W} \times \mathrm{W}$ es una relación de equivalencia entre mundos, y
- $\mathrm{V} : \mathrm{W} \to 2^{\mathsf{Prop}}$ es una función de valuación que a cada mundo le asigna el conjunto de proposiciones que se cumplen en dicho mundo.

Dado que las relaciones $\sim_i$ son simétricas, usualmente los modelos se los representa gráficamente como un grafo no dirigido, donde a cada arista se la etiqueta con la correspondiente relación $\sim_i$ asociada, omitiendo las aristas reflexivas. Notar que una arista puede ser compartida por múltiples agentes.

Ejemplo 2.1. Retomemos el Ejemplo 1.5. Sean dos agentes i y j tales que se encuentran dentro de una casa. El agente i se encuentra lejos de cualquier ventana o abertura. Por ende, desde su perspectiva, no distingue si el día está soleado o nublado. En cambio, el agente j se encuentra al lado de la ventana. En cualquiera de las dos situaciones, si está nublado o soleado, el agente podrá distinguir ambos mundos entre sí. Un modelo que representa esta situación es $\mathcal{M} = \langle \mathrm{W}, \{\sim_i, \sim_j\}, \mathrm{V}\rangle$ tal que

- $\mathrm{W} = \{w_1, w_2\}$,
- $\sim_i = \{(w_1, w_1), (w_2, w_2), (w_1, w_2), (w_1, w_2)\}$,
- $\sim_j = \{(w_1, w_1), (w_2, w_2)\}$, y
- $\mathrm{V}(w_1) = \{s\}$, $\mathrm{V}(w_2) = \{n\}$;

siendo s y n las variables proposicionales que representan las situaciones *el día está soleado* y *el día está nublado* respectivamente. Su representación gráfica es la siguiente:



$\square$

La semántica de K_i se define de manera análoga a la de $\Box$, esta vez sobre la relación $\sim_i$ correspondiente. Considerando la semántica de la Definición 1.3, dado un modelo $\mathcal{M} = \langle \mathrm{W}, \{\sim_i\}_{i\in\mathsf{Agt}}, \mathrm{V}\rangle$ y un mundo $w \in \mathrm{W}$, un agente i sabe que una proposición φ es verdadera ($\mathsf{K}_i\varphi$) en w si y solamente si φ se cumple en todos los mundos w' que el agente no puede distinguir de w. Es decir,

$$\mathcal{M}, w \models \mathsf{K}_i\varphi \text{ sii para todo } w' \in \mathrm{W} \text{ tal que } w \sim_i w', \text{ se tiene que } \mathcal{M}, w' \models \varphi.$$

Axiomas:	TAUT	$\vdash \varphi$ para φ una tautología proposicional
	DIST	$\vdash \mathsf{K}_i(\varphi \to \psi) \to (\mathsf{K}_i\varphi \to \mathsf{K}_i\psi)$
	T	$\vdash \mathsf{K}_i\varphi \to \varphi$
	4	$\vdash \mathsf{K}_i\varphi \to \mathsf{K}_i\mathsf{K}_i\varphi$
	5	$\vdash \widehat{\mathsf{K}}_i\varphi \to \mathsf{K}_i\widehat{\mathsf{K}}_i\varphi$
Reglas:	$\dfrac{\varphi \quad (\varphi \to \psi)}{\psi}$ MP	$\dfrac{\varphi}{\mathsf{K}_i\varphi}$ NEC

Tabla 2.1: Sistema axiomático de $\mathsf{L_K}$

Ejemplo 2.2. Teniendo en cuenta el modelo del Ejemplo 2.1, $\mathcal{M}, w_1 \not\models \mathsf{K}_i s$, ya que el agente i no distingue w_1 de w_2 ($w_1 \sim_i w_2$) y este último no cumple s. Sin embargo, en el caso del agente j, $\mathcal{M}, w_1 \models \mathsf{K}_j s$ ya que el agente puede distinguir w_1 de los demás. Análogamente, $\mathcal{M}, w_2 \models \mathsf{K}_j n$. □

En la Tabla 2.1 se introduce un sistema axiomático de esta lógica. El mismo establece que es una *lógica modal normal*, dado que cada una de las modalidades K_i satisface los axiomas y reglas básicas de la Tabla 1.1. El resto de los axiomas (T, 4 y 5) se encargan de describir que cada $\sim_i$ es una relación de equivalencia. La Tabla 2.1 es correcta y fuertemente completa con respecto a la clase de modelos donde cada una de las relaciones $\sim_i$ es una relación de equivalencia [24, 26].

Teorema 2.1 ([24,26])**.** *El sistema axiomático de la Tabla 2.1 es correcto y fuertemente completo con respecto a la clase de todos los modelos de Kripke epistémicos multiagente.*

Sin embargo, aún teniendo estos modelos, uno tiene una visión estática del conocimiento. Los agentes saben sobre ciertas proposiciones, pero no son capaces de cambiar su conocimiento. Consideremos el Ejemplo 2.1 sobre los agentes encerrados en una habitación sin visión del mundo exterior. Naturalmente, si el agente que no tiene ninguna idea acerca del estado del tiempo abre una de sus ventanas, podrá adquirir nueva información, cambiando su noción de incertidumbre y así el conocimiento acerca de ciertos hechos. Resulta natural entonces contar con maneras lógicas de representar la información que cambia mediante la aplicación de ciertas operaciones o acciones. Para estos casos, se consideran operadores dinámicos (ver por ejemplo, [7]) que actualizan diferentes partes del modelo.

Consideremos el operador dinámico de anuncios públicos [91, 109], el cual informa a todos los agentes que una determinada fórmula χ es verdadera, denotada como $[\chi]$. Una fórmula $[\chi]\varphi$ se lee como *"luego de que se anuncia públicamente que χ es verdadera, vale φ"*. Esta acción se traduce semánticamente como una modificación del modelo mediante la eliminación de estados (e indirectamente relaciones) que ya no cumplan con esta información. De esta manera, el agente ya no considera posible ciertas situaciones, y por ende disminuye la incertidumbre sobre los hechos. A este tipo de anuncios se los denomina "públicos", ya que la información es revelada a todos los agentes por igual, es decir, el anuncio afecta a la incertidumbre de todos los agentes.

La semántica de este operador en términos de modelos de Kripke es la siguiente:

$$\mathcal{M}, w \models [\chi]\varphi \;\text{ sii }\; \mathcal{M}, w \models \chi, \text{ entonces } \mathcal{M}_\chi, w \models \varphi,$$

con $\mathcal{M}_\chi = \langle \mathrm{W}_\chi, \mathrm{R}_\chi, \mathrm{V}_\chi \rangle$ donde

- $\mathrm{W}_\chi = \{w \mid \mathcal{M}, w \models \chi\} = [\![\chi]\!]^{\mathcal{M}}$,
- $\mathrm{R}_\chi = \mathrm{R} \cap \mathrm{W}_\chi^2$, y
- $\mathrm{V}_\chi(w) = \mathrm{V}(w)$ para todo $w \in \mathrm{W}_\chi$.

En otras palabras, el operador elimina los mundos que no cumplen χ y restringe las relaciones y la función de valuación al nuevo dominio de $[\![\chi]\!]^{\mathcal{M}}$. Considerando el modelo del Ejemplo 2.1, si bien $\mathcal{M}, w_1 \not\models \mathsf{K}_i s$, si se anuncia a todos los agentes que está soleado ($[s]$), el agente i ya no considera posible el mundo nublado w_2 y por lo tanto sabe que el día está soleado. Es decir, $\mathcal{M}, w_1 \models [s]\mathsf{K}_i s$.

Retomando la Definición 1.6 y las discusiones posteriores alrededor de los Ejemplos 1.11 y 1.12, cuando se extiende una lógica con operadores dinámicos, surge la pregunta acerca de si

RAtom	$\vdash [\chi]p \leftrightarrow (\chi \to p)$
$\mathsf{R}\neg$	$\vdash [\chi]\neg\varphi \leftrightarrow (\chi \to \neg[\chi]\varphi)$
$\mathsf{R}\vee$	$\vdash [\chi](\varphi \vee \psi) \leftrightarrow [\chi]\varphi \vee [\chi]\psi$
RK	$\vdash [\chi]\mathsf{K}_i\varphi \leftrightarrow (\chi \to \mathsf{K}_i[\chi]\varphi)$
$\mathsf{RE}_{[\chi]}$	$\vdash [\chi][\chi']\varphi \leftrightarrow [\chi \wedge [\chi]\chi']\varphi$

Tabla 2.2: Axiomas de reducción de PAL

es posible expresar las nuevas modalidades en el lenguaje base. En otras palabras, si la lógica original es lo suficientemente expresiva para poder emularlas.

Consideremos la extensión de $\mathsf{L_K}$ con un operador dinámico, denotada como $\mathsf{L_K^+}$. Naturalmente, $\mathsf{L_K} \preceq \mathsf{L_K^+}$ tomando la función de traducción identidad, es decir, $T : \mathsf{L_K} \to \mathsf{L_K^+}$ tal que $T(\varphi) = \varphi$. Para demostrar que $\mathsf{L_K^+}$ es reducible a $\mathsf{L_K}$, es decir que $\mathsf{L_K^+} \preceq \mathsf{L_K}$, una estrategia usual en lógica modal es la definición de axiomas de reducción (ver [106, 109] para más detalles). Los axiomas de reducción son una serie de equivalencias entre fórmulas de la lógica tal que, vistas como una reescritura de una fórmula en otra, nos permiten eliminar todas las ocurrencias de un operador. En el caso de $\mathsf{L_K^+}$, la idea es definir axiomas de reducción tal que el operador dinámico pueda ser eliminado. Los axiomas de reducción además, en conjunto con la axiomatización para $\mathsf{L_K}$, nos permite obtener una axiomatización para $\mathsf{L_K^+}$.

Como mencionamos entonces obtenemos una traducción, codificada en estos axiomas que toma una fórmula en el lenguaje 'dinámico' y la lleva a una equivalente en el lenguaje 'base'. Por supuesto, esta estrategia depende de la riqueza del lenguaje de $\mathsf{L_K}$. La existencia de estos axiomas de reducción demuestran que el lenguaje base es lo suficientemente expresivo para describir los cambios en el modelo que la modalidad dinámica induce [106].

Ejemplo 2.3. Sea PAL la extensión de $\mathsf{L_K}$ con $[\chi]$. Como se puede apreciar por ejemplo en [109], contamos con axiomas de reducción (Tabla 2.2) que son válidos en todos los modelos. Basándonos en estos, se puede definir inductivamente la siguiente función de traducción $T : \mathsf{PAL} \to \mathsf{L_K}$:

- $T(p) = p$
- $T(\neg\varphi) = \neg T(\varphi)$,
- $T(\varphi \vee \varphi) = T(\varphi) \vee T(\varphi)$,
- $T([\chi]p) = (T(\chi) \to p)$,
- $T([\chi]\neg\varphi) = (T(\chi) \to \neg T([\chi]\varphi))$,
- $T([\chi](\varphi \vee \psi)) = T([\chi]\varphi) \vee T([\chi]\psi)$,
- $T([\chi]\mathsf{K}_i\varphi) = (T(\chi) \to \mathsf{K}_i T([\chi]\varphi))$, y
- $T([\chi][\chi']\varphi) = [T(\chi) \wedge T([\chi]\chi')]T(\varphi)$.

Con esto, $\mathsf{PAL} \preceq \mathsf{L_K}$ y $\mathsf{PAL} \equiv \mathsf{L_K}$. □

Notar que la traducción reduce el tamaño de la fórmula, luego es posible eliminar las ocurrencias de $[\chi]$, y por ende siempre termina. Más aún, dado que $\mathsf{L_K}$ ya posee una axiomatización completa (Teorema 2.1), junto con estos axiomas de reducción se tendría completitud para $\mathsf{L_K^+}$, como lo es en el caso de PAL.

Teorema 2.2 ([26, 109])**.** *El sistema axiomático de* PAL*, compuesto por las Tablas 2.1 y 2.2, es correcto y fuertemente completo con respecto a la clase de todos los modelos.*

A pesar de su simplicidad, esta representación del conocimiento basado en la indistinguibilidad tiene varias ventajas. En primer lugar, se captura el conocimiento de alto orden del agente (conocimiento sobre el propio conocimiento o el de otros agentes). En segundo lugar, debido a su generalidad, se pueden estudiar otros conceptos epistémicos como la *creencia* [62], el conocimiento colectivo [38, 39], entre otros. Y finalmente, permite una representación natural de acciones a través de las cuales el conocimiento cambia a partir de la definición de operadores dinámicos [106, 109]. Esta manera de representar el conocimiento será nuestra guía a la hora de modelar otros tipos de conocimento, con otras particularidades.

2.2. Otros patrones de conocimento

Las lógicas epistémicas de *saber qué* [62] han influenciado considerablemente a las ciencias de la computación. Sin embargo, existen otros patrones de conocimiento que han sido estudiados en menor medida [117]. Por ejemplo, si quisiéramos expresar que un agente sabe acerca de un evento en particular (el resultado de un partido de fútbol o de una elección), uno hablaría de una noción de *saber si* (el agente sabe si ganó tal equipo o si ganó tal candidato). Ejemplos de estas lógicas están en [40,56,117]. Otro tipo de conocimento estudiado es el de describir que un agente sabe las razones de una situación (por ejemplo, las acciones de un árbitro o un político). El mismo describe el concepto de *saber por qué* (el agente sabe por qué el árbitro ha solicitado el VAR o sabe por qué un político ha aplicado una determinada medida). Particularmente en [13,119] se introducen lógicas de justificaciones, orientadas a caracterizar formalmente el conocimiento de las razones y/o explicaciones de ciertos escenarios. Por otro lado, una alternativa más abarcativa es la idea de *saber el valor* que generaliza los objetos de los saberes con términos y variables [16, 52, 111]. Específicamente, en esta tesis se enfocará en un patrón de conocimiento particular, motivado por los campos de la filosofía y la inteligencia artificial, el cual se enfoca en las habilidades del agente: la noción de *saber cómo* [41]. En general se considera que un agente sabe cómo alcanzar φ dado ψ si tiene la habilidad de garantizar que φ se cumple desde una situación en la que ψ se cumple. El significado de tener una habilidad que garantice el objetivo es uno de los objetos de estudio y discusión.

Históricamente, el concepto de "saber cómo" ha sido considerado diferente del "saber qué" [89, 94]. El saber cómo puede verse como un reflejo de los cursos de acción o habilidades que el agente puede tomar para cumplir un objetivo. Existe una vasta literatura que conecta el *saber cómo* con las lógicas de conocimento y acción [61, 71, 80, 83, 107]. Sin embargo, estas representaciones del *saber cómo* han sido el objeto de diversas críticas, argumentando que la combinación de los operadores que expresan *saber qué* y *habilidades* no lleva a una noción natural de *saber cómo* [59,66]. En general, en estos enfoques, el saber cómo suele modelarse con expresiones de la forma $\mathsf{K}\Diamond\varphi$, interpretada como *"el agente sabe que existe un curso de acción que lo/la lleve a un estado donde vale φ"*. Esto es lo que se conoce como una lectura *de dicto* del saber cómo. Esto difiere de la lectura *de re* del conocimiento [103], en el cual el patrón de cuantificación se invierte (existe un método tal que el agente sabe que es el apropiado para lograr el objetivo), la cual en general se considera más adecuada. Más aún, en la literatura se argumenta que el operador debe ser global, es decir, que no dependa de la situación actual en la cual se encuentra el agente.

Considerando esto, en [116–118] se introduce un nuevo marco teórico basado en una modalidad de *saber cómo* binaria que no está definida en términos de *saber qué*. A nivel semántico, esta lógica se interpreta sobre modelos relacionales llamados sistemas de transición etiquetadas (LTSs). Los LTSs son esencialmente modelos de Kripke, donde los mundos representan estados y las relaciones de accesibilidad representan la ejecución de ciertas acciones, algo similar a lo que ocurre en la lógica dinámica proposicional (PDL) [26, 45, 55, 104] considerada en el capítulo anterior (Ejemplo 1.3). En un LTS, un arco etiquetado como a de un estado w a uno u indica que el agente puede ejecutar la acción a para transformar el estado w en u. En la semántica propuesta, la nueva modalidad denotada como $\mathsf{Kh}(\psi, \varphi)$ se cumple si y solamente si existe un "plan", es decir una secuencia de acciones, que cumpla una condición llamada ejecutabilidad fuerte que a partir de cada estado que cumpla la "precondición" ψ, luego de ejecutar el plan nos garantice llegar a estados donde se cumpla la "postcondición" φ. Esta condición de ejecutabilidad fuerte implica que el plan sea "a prueba de fallas" en el LTS. Es decir que, de manera infalible, cada vez que el plan comienza a ejecutarse, su ejecución puede ser completada. Esto se justifica ya que la ejecución de acciones es no determinística. La condición de ejecutabilidad fuerte se hereda del campo de *planning* [49]. Existen otras variantes de *saber cómo* que siguen un enfoque similar [42, 73, 76, 115].

Algo importante a destacar de esta propuesta es que los LTSs no tienen una componente epistémica. Dado que las relaciones son interpretadas como acciones, las habilidades de un agente sólo están definidas a partir de qué pueden conseguir con estas acciones. Comparándolo con la lógica epistémica de *saber qué* discutida anteriormente, los modelos de esta proveen dos tipos de información: ***(1)*** información denominada *óntica* o factual, la cual es provista por el mundo donde se evalúa la fórmula y los símbolos de proposición que valen en él (W y V) y ***(2)*** la información *epistémica* provista por la relación de indistinguibilidad para cada agente ($\{\sim_i\}_{i \in \mathsf{Agt}}$). Esta última representa las situaciones que el agente "considera" como posibles. En un contexto multiagente, todos los agentes comparten la misma información óntica pero difieren en su *interpretación epistémica*.

Esta separación deja la pregunta abierta de si resulta posible definir el *saber cómo* utilizando una noción de indistinguibilidad, o al menos de hacer esta distinción entre información óntica y epistémica, de tal forma que los modelos sean capaces de capturar las habilidades del agente dadas por las acciones disponibles (la información óntica) y el conocimiento (o falta de él) representada por su incertidumbre entre diferentes cursos de acción (la información epistémica).

Teniendo en cuenta esto, en este trabajo se investiga una nueva semántica para $\mathsf{Kh}_i(\psi, \varphi)$, una versión multiagente de la modalidad de *saber cómo* presentada en [8]. Dicha propuesta permite usar una noción de *indistinguibilidad entre planes*, en vez de la indistinguibilidad entre mundos de la lógica epistémica clásica. De esta manera separamos la información óntica brindada por las habilidades o cursos de acción disponibles por la estructura del LTS, de la información epistémica en este caso descripta por una *relación de indistinguibilidad entre planes*. Esta idea está basada en la noción de *indistinguibilidad entre estrategias*, introducida por ejemplo en [21, 67]. Esto permite considerar otras razones por las cuales un agente no sepa cómo alcanzar un objetivo, aparte de que no hayan acciones disponibles en el LTS. Por ejemplo, el agente tiene un plan adecuado pero no es capaz de distinguirlo de otro que no lo es. Es decir, no es consciente que ambos planes producen diferentes resultados. Otra razón es que existe un plan disponible *ontológicamente* (está presente en el LTS) pero no *epistémicamente* para un determinado agente. Dicho agente puede no ser *consciente* de todos los cursos de acción posibles y sólo considere un conjunto limitado de estos. Esta idea de consciencia en lógicas epistémicas se ha usado para lidiar con los problemas de omnisciencia lógica [54, 102, 112], permitiendo que el agente no sea consciente de todas las fórmulas involucradas. Con esto se consigue modelar agentes más "reales" y con ciertas limitaciones [38]. En otras palabras, el saber cómo de un agente implica la existencia de su habilidad (es decir, su capacidad de hacerlo), pero tener dicha habilidad no implica necesariamente que sepa cómo [57], a diferencia de los enfoques previos que contaban con un cierto nivel e omnisciencia.

2.3. Contenido de esta tesis

Este trabajo está conformado por un compendio de investigaciones realizadas a lo largo del doctorado en [5,6,8–10]. En el Capítulo 3 se introduce una lógica de saber cómo para la modalidad $\mathsf{Kh}(\psi, \varphi)$ interpretada sobre planes lineales y en sistemas de transiciones etiquetadas (LTSs) [116–118]. Se darán sus propiedades de axiomatización, bisimulaciones, equivalencia de fórmulas, complejidad computacional del problema de model checking [36] y el problema de satisfacibilidad, siendo este último un problema abierto resuelto en este doctorado [6]. Además, discutiremos las limitaciones de esta lógica y mostraremos otras variantes del operador Kh [42, 72, 73, 76, 115].

En el Capítulo 4 se define una lógica de saber cómo multiagente con una nueva semántica para la modalidad $\mathsf{Kh}_i(\psi, \varphi)$ (para un agente i) en la cual se reintroduce la noción de indistinguibilidad epistémica de la lógica epistémica clásica. Esta nueva dimensión captura, para cada agente particular, la conciencia de las habilidades disponibles en el mundo real. Esto incorpora situaciones de incertidumbre no contempladas en el capítulo anterior. En el Capítulo 5 se presenta una axiomatización para la lógica acompañada de sus correspondientes pruebas de correctitud y completitud fuerte. En el Capítulo 6 se obtienen resultados de bisimulaciones (teoremas de invarianza y Hennessy-Milner para modelos finitos) usando una variación de la definición de bisimulaciones de [43, 44]. Además, se prueba que la lógica es más débil que la propuesta original [116–118] pero a la vez la generaliza. En el Capítulo 7 se demuestra que el problema de satisfacibilidad es decidible vía filtraciones para modelos arbitrarios y es NP-completo mediante un método de selección en el modelo canónico. Todos estos resultados han sido demostrados en [8] y en su versión extendida en [10].

En el Capítulo 8, se introducen operadores dinámicos que modifican la información óntica (basados en anuncios públicos [91] y actualización de relaciones [69]) y epistémica (refinando la relación de indistinguibilidad) de los modelos. Para los primeros se tienen axiomatizaciones y resultados de decidibilidad sobre clases de modelos restringidas. Para los segundos se presentan diferentes propuestas para representar el aprendizaje de los agentes y las dificultades de conseguir axiomatizaciones para las lógicas resultantes. Estas propiedades fueron presentadas en [9]. Con estas nociones, en el Capítulo 9 se enriquece el lenguaje base con modalidades que representan ejecuciones de acciones explícitas, y se presentan operadores dinámicos con una axiomatización completa con respecto a la clase de todos los modelos. Las demostraciones para estas últimas propiedades se muestran por primera vez en este trabajo.

Por otro lado, en el Capítulo 10, tomando inspiración de [23], se reinterpreta la modalidad de saber cómo para introducir una lógica deóntica que considera tres modalidades, así como se realizó en [5], que cubren tres aspectos condicionantes para un agente: ***(1)*** sus habilidades ($\mathsf{S}(\psi, \varphi)$), ***(2)*** las normas de su entorno ($\mathsf{N}(\psi, \varphi)$) y ***(3)*** su nivel de consciencia sobre sus responsabilidades ($\mathsf{Kc}_i(\psi, \varphi)$). En el mismo, se presentan resultados de complejidad vía una función de selección para las fórmulas con los operadores N y Kc_i y dos axiomatizaciones, una para N y Kc_i y otra para las tres modalidades juntas. Finalmente, en los Capítulos 11 y 12 se hablan de las conclusiones y el trabajo futuro, respectivamente.

Parte II

Saber cómo en Lógica Epistémica

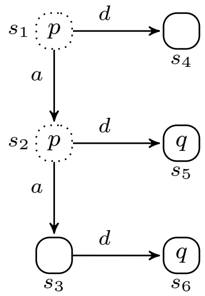
Capítulo 3

Una lógica de saber cómo sobre planes lineales

> A centipede was happy – quite!
> Until a toad in fun
> Said, "Pray, which leg moves after which?"
> This raised her doubts to such a pitch,
> She fell exhausted in the ditch
> Not knowing how to run.
>
> The Centipede's Dilemma

En este capítulo se presentará uno de los enfoques para modelar el concepto de saber cómo, desde un punto de vista lógico, que tiene mayor consenso en el área: la lógica de saber cómo introducida en [116–118], interpretada sobre *planes lineales*. En este enfoque las habilidades de un agente están representadas por Sistemas de Transiciones Etiquetadas (LTS por sus siglas en inglés), donde las transiciones etiquetadas por símbolos de acción reflejan los efectos de ejecutar una acción en ciertos estados. A su vez, la particularidad de este enfoque es que dichas habilidades reflejan directamente el conocimiento del agente. Es decir, que de alguna manera las componentes ónticas y epistémicas coinciden (este aspecto será discutido en profundidad más adelante). Las habilidades y el conocimiento del agente serán obtenidos no sólo a partir de las acciones atómicas y sus efectos, sino que se considerará la composición de las mismas. Dicha composición de acciones se llamarán *planes*. El siguiente ejemplo, extraído de [116], nos ayudará a entender la manera en la que las habilidades de un agente son modeladas en un LTS.

Ejemplo 3.1. Consideremos que queremos representar el mapa de un edificio. En un LTS, cada ubicación del edificio puede ser modelada como un estado. Si representamos gráficamente al LTS como un grafo dirigido, los estados se corresponden con los nodos del mismo y pueden estar etiquetados con símbolos de proposición que indican que cierta propiedad vale en el estado. Por ejemplo, en el LTS a continuación, hay ciertos estados etiquetados con el símbolo p (s_1 y s_2), otros con q (s_5 y s_6), y otros donde ninguno de ellos ocurre (s_3 y s_4). El símbolo q indica que la ubicación es un lugar seguro en el edificio. En esta representación existen dos acciones: moverse a la derecha (d) y moverse hacia abajo (a). El agente, por ejemplo, al moverse desde s_1 hacia la derecha, alcanza el estado s_4, lo cual se representa con una transición etiquetada por d, que va desde s_1 a s_4. Además, el agente puede ejecutar planes, es decir, la composición de acciones básicas, tales como ad (primero abajo y luego a la derecha) ó aa (moverse dos veces abajo). En particular, en cada estado donde vale p, existe un plan (ad) que le permite al agente alcanzar un lugar seguro (q). En cambio, el plan aa en ocasiones falla a la hora de alcanzar un lugar seguro. Por ejemplo, al ser ejecutado en s_2, el agente puede moverse una vez hacia abajo, pero no puede hacerlo una segunda vez desde s_3. Por otro lado, en s_1 es posible ejecutar aa, pero se alcanza un estado (s_3) donde no vale q (es decir, no es un lugar seguro).



□

A continuación se introducirán los fundamentos y las nociones elementales de la lógica. Se definirán la sintaxis de las fórmulas y la semántica de las mismas, que se interpretan sobre estos LTSs. Además, se presentará un sistema axiomático: Un conjunto de axiomas, fórmulas que se cumplen en cualquier modelo, y de reglas, silogismos formados por una serie de premisas y una conclusión. Usando estos elementos se pueden deducir otras fórmulas llamadas teoremas. De manera más general, el sistema axiomático presenta dos propiedades importantes. La primera es que es fuertemente completo, es decir que si una fórmula φ es consecuencia lógica de un conjunto de fórmulas Γ (denotado como $\Gamma \models \varphi$) entonces φ se deduce de Γ en el sistema (denotado como $\Gamma \vdash \varphi$). La segunda, que es la recíproca de fuertemente completo, es que es correcto, y se reduce a demostrar que las reglas y axiomas se cumplen en todos los modelos. También se presentarán conceptos estándar en lógicas modales como resultados relacionados a bisimulaciones, con sus teoremas de invarianza y Hennessy-Milner, y complejidad computacional en torno a los problemas de model checking y satisfacibilidad. En cada una de estas secciones se discutirán algunas de estas propiedades. Finalmente, se comentarán sobre otras lógicas de saber cómo alternativas a la vista en profundidad.

3.1. Sintaxis y semántica

En esta sección se introducirán la sintaxis y la semántica de la lógica de saber cómo, denotada como $\mathsf{L_{Kh}}$ de [116–118]. A lo largo de esta tesis, se define Prop como un conjunto numerable no vacío de símbolos proposicionales.

Definición 3.1. Las fórmulas de la lógica $\mathsf{L_{Kh}}$ son dadas por la gramática

$$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid \mathsf{Kh}(\varphi, \varphi),$$

con $p \in \mathsf{Prop}$. Las constantes y demás operadores booleanos se definen como:

- $\top := (p \vee \neg p)$,
- $\bot := \neg\top$,
- $(\varphi \to \psi) := (\neg\varphi \vee \psi)$,
- $(\varphi \wedge \psi) := \neg(\neg\varphi \vee \neg\psi)$,
- $(\varphi \leftrightarrow \psi) := (\varphi \to \psi) \wedge (\psi \to \varphi)$.

Las fórmulas de la forma $\mathsf{Kh}(\psi, \varphi)$ se leen como *“cuando ψ se cumple, el agente sabe cómo hacer φ verdadera”*. En algunos casos, a ψ se la refiere como “precondición” y mientras que a φ como “postcondición”.

En [116–118] (y variaciones como [73, 76]), las formulas de $\mathsf{L_{Kh}}$ son interpretadas sobre *sistemas de transiciones etiquetadas*: modelos relacionales en los cuales las relaciones describen las transiciones a otros estados disponibles para el agente. Dichas relaciones se definen a partir de acciones básicas o atómicas que, concatenadas con otras, constituyen planes.

Definición 3.2 (Acciones y planes)**.** Sea Act un conjunto numerable de nombres de acciones (acciones básicas), se denota con Act^* al conjunto de secuencias finitas sobre Act. Los elementos de Act^* son llamados *planes*, con ϵ siendo el *plan vacío*.

Ejemplo 3.2. Supongamos que se tiene un agente en un tablero de ajedrez, y que este puede moverse en cuatro direcciones posibles: arriba ($\uparrow$), abajo ($\downarrow$), izquierda ($\leftarrow$) y derecha ($\rightarrow$). En este contexto, el conjunto de acciones básicas a considerar es $\mathsf{Act} = \{\uparrow, \downarrow, \leftarrow, \rightarrow\}$ y Act^* es el conjunto de todos los movimientos que puede hacer este agente (incluso el de no hacer ningún movimiento, el plan ϵ). Por ejemplo, $\uparrow\uparrow\leftarrow$ y $\rightarrow\downarrow\leftarrow\downarrow$ están en Act^*. $\square$

Definición 3.3. Sea $\sigma \in \mathsf{Act}^*$, se denota $|\sigma|$ a la longitud de σ, con $|\epsilon| := 0$. Dado un plan σ y $0 \leq k \leq |\sigma|$, el *plan* σ_k es el segmento inicial de σ hasta la k-ésima posición (con $\sigma_0 := \epsilon$). Para $0 < k \leq |\sigma|$, la *acción* $\sigma[k]$ es la que se encuentra en la k-ésima posición de σ.

Ejemplo 3.3. Tomando el conjunto Act del Ejemplo 3.2, $|\rightarrow\downarrow\leftarrow\downarrow| = 4$ y se tiene que:

- $(\rightarrow\downarrow\leftarrow\downarrow)_0 = \epsilon$,
- $(\rightarrow\downarrow\leftarrow\downarrow)_1 = \rightarrow$,
- $(\rightarrow\downarrow\leftarrow\downarrow)_2 = \rightarrow\downarrow$,
- $(\rightarrow\downarrow\leftarrow\downarrow)_3 = \rightarrow\downarrow\leftarrow$, y
- $(\rightarrow\downarrow\leftarrow\downarrow)_4 = \rightarrow\downarrow\leftarrow\downarrow$.

Además, $(\rightarrow\downarrow\leftarrow\downarrow)[1] = \rightarrow$, $(\rightarrow\downarrow\leftarrow\downarrow)[2] = \downarrow$, $(\rightarrow\downarrow\leftarrow\downarrow)[3] = \leftarrow$ y $(\rightarrow\downarrow\leftarrow\downarrow)[4] = \downarrow$. $\square$

A continuación se introduce la definición de los modelos de la lógica L_{Kh}.

Definición 3.4 (Sistema de transiciones etiquetadas)**.** Un *sistema de transiciones etiquetadas (*LTS*)* para Prop y $A \subseteq \mathsf{Act}$ es una tupla $\mathfrak{L} = \langle \mathrm{S}, \mathrm{R}, \mathrm{V} \rangle$, donde:

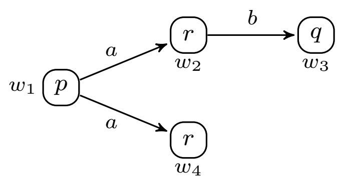
- S es un conjunto no vacío de estados (también denotado como $\mathrm{D}_{\mathfrak{L}}$),
- $\mathrm{R} = \{\mathrm{R}_a \subseteq \mathrm{S} \times \mathrm{S} \mid a \in A\}$ es una colección de relaciones binarias sobre S, y
- $\mathrm{V} : \mathrm{S} \to 2^{\mathsf{Prop}}$ es una función de etiquetado.

Un LTS $\mathfrak{L} = \langle \mathrm{S}, \mathrm{R}, \mathrm{V} \rangle$ se representa gráficamente como un grafo dirigido, en el cual los nodos son los estados de S, cada uno marcado con las variables proposicionales determinadas por la función de etiquetado V, y las aristas entre dos estados son los pares definidos en R_a por cada acción básica $a \in A$.

Ejemplo 3.4. Sea $\mathfrak{L} = \langle \mathrm{S}, \mathrm{R}, \mathrm{V} \rangle$ un LTS tal que

- $\mathrm{S} = \{w_1, w_2, w_3, w_4\}$;
- $\mathrm{R} = \{\mathrm{R}_a, \mathrm{R}_b\}$ con $\mathrm{R}_a = \{(w_1, w_2), (w_1, w_4)\}$ y $\mathrm{R}_b = \{(w_2, w_3)\}$; y
- $\mathrm{V}(w_1) = \{p\}$, $\mathrm{V}(w_2) = \mathrm{V}(w_4) = \{r\}$, $\mathrm{V}(w_4) = \{q\}$.

Su representación gráfica es la siguiente:



$\square$

En términos generales, un LTS describe las *habilidades* de un agente por lo que a veces (por ejemplo, en [116–118]) es llamado un *mapa de habilidades*. En las siguientes definiciones se introducen algunos conceptos necesarios sobre dichas habilidades. Si bien el conjunto Act puede ser infinito, las relaciones en el modelo pueden estar definidas para un subconjunto (posiblemente finito) de acciones.

Definición 3.5. Sean $A \subseteq \mathsf{Act}$ y $\{\mathrm{R}_a \subseteq \mathrm{S} \times \mathrm{S} \mid a \in A\}$ una colección de relaciones binarias sobre S. Se definen $\mathrm{R}_\epsilon := \{(w, w) \mid w \in \mathrm{S}\}$ y, para $\sigma \in A^*$ y $a \in A$,

$$\mathrm{R}_{\sigma a} := \{(w, u) \in \mathrm{S} \times \mathrm{S} \mid \text{existe } v \in \mathrm{S} \text{ tal que } (w, v) \in \mathrm{R}_\sigma \text{ y } (v, u) \in \mathrm{R}_a\}.$$

Dado un plan $\sigma \in A^*$, para $u \in \mathrm{S}$, $\mathrm{R}_\sigma(u) := \{v \in \mathrm{S} \mid (u, v) \in \mathrm{R}_\sigma\}$, y para $U \subseteq \mathrm{S}$, $\mathrm{R}_\sigma(U) := \bigcup_{u \in U} \mathrm{R}_\sigma(u)$.

Ejemplo 3.5. Usando el modelo del Ejemplo 3.4, $(w_1, w_3) \in \mathrm{R}_{ab}$ ya que existe $v = w_2 \in \mathrm{S}$ tal que $(w_1, v) \in \mathrm{R}_a$ y $(v, w_3) \in \mathrm{R}_b$. Como es el único par que cumple esto, $\mathrm{R}_{ab}(w_1) = \{w_3\}$. □

La idea en [116–118] es que el agente sabe cómo conseguir φ dado ψ cuando tiene un plan apropiado que le permite ir desde cualquier estado el cual se cumple ψ, sólo a estados donde φ se cumple. La parte importante es determinar cuándo un plan es "apropiado". En este caso, los planes apropiados son aquellos llamados "fuertemente ejecutables" en los estados que cumplan ψ. Esto es que, una vez que empieza a ejecutarse dicho plan desde esos estados, todas las posibles ejecuciones parciales siempre terminan. La noción proviene del área de *conformant planning* [100] para planes lineales en las que se considera el no determinismo a la hora de ejecutar ciertos cursos de acción. La realización de una misma acción o plan desde un punto determinado puede llevar a diferentes estados en diferentes ejecuciones.

Definición 3.6 (Ejecutabilidad fuerte)**.** Sea $\{\mathrm{R}_a \subseteq \mathrm{S} \times \mathrm{S} \mid a \in A\}$ una colección de relaciones binarias. Un plan $\sigma \in \mathsf{Act}^*$ es *fuertemente ejecutable* (SE) en $u \in \mathrm{S}$ si y sólo $\sigma \in A^*$ y para cada $k \in [0 \,...\, |\sigma| - 1]$, $v \in \mathrm{R}_{\sigma_k}(u)$ implica $\mathrm{R}_{\sigma[k+1]}(v) \neq \emptyset$. Con esto se define $\mathrm{SE}(\sigma) := \{w \in \mathrm{S} \mid \sigma \text{ es fuertemente ejecutable en } w\}$.

Ejemplo 3.6. Teniendo como referencia el Ejemplo 3.4, para $\sigma = \epsilon$, como $|\epsilon| = 0$, se tiene que $[0 \,...\, |\sigma| - 1] = [0 \,...\, -1] = \emptyset$. Por rango vacío, ϵ es fuertemente ejecutable en todo estado. Por lo tanto, $\mathrm{SE}(\epsilon) = \mathrm{S}$.

Para el caso $\sigma = a$, $|a| = 1$ y $[0 \,...\, |\sigma| - 1] = [0 \,...\, 0] = \{0\}$. Por lo tanto, basta con ver que $v \in \mathrm{R}_{\sigma_k}(u)$ implica $\mathrm{R}_{\sigma[k+1]}(v) \neq \emptyset$ para $k = 0$. Esto es equivalente a $v \in \mathrm{R}_\epsilon(u)$ implica $\mathrm{R}_a(v) \neq \emptyset$ (pues $\sigma_0 = \epsilon$ y $\sigma[1] = a$). Sea $v \in \mathrm{R}_\epsilon(u)$ $(u = v)$, esto implica $\mathrm{R}_a(v) \neq \emptyset$ si y sólo si $v = w_1$. Con esto, a es fuertemente ejecutable únicamente en w_1 y $\mathrm{SE}(a) = \{w_1\}$. Con un razonamiento similar, $\mathrm{SE}(b) = \{w_2\}$. Notar que para toda acción atómica $c \in \mathsf{Act}$, $\mathrm{R}_c(s) \neq \emptyset$ si y solamente si c es fuertemente ejecutable en s.

Para $\sigma = ab$, $[0 \,...\, |\sigma| - 1] = [0 \,...\, 1] = \{0, 1\}$. El único estado considerable es w_1 pues para $k = 0$, es el único que cumple $v \in \mathrm{R}_\epsilon(w_1)$ implica $\mathrm{R}_a(v) \neq \emptyset$. Sin embargo, para $k = 1$ $(\sigma_k = a$, $\sigma[k+1] = b)$, $v \in \mathrm{R}_a(w_1)$ implica $\mathrm{R}_b(v) \neq \emptyset$ no siempre se cumple. Dado que $\mathrm{R}_a(w_1) = \{w_2, w_4\}$, $\mathrm{R}_b(w_2) = \{w_3\} \neq \emptyset$ pero $\mathrm{R}_b(w_4) = \emptyset$. Por lo tanto, $\mathrm{SE}(ab) = \emptyset$. Se puede ver que para cualquier $\sigma \in \{a, b\}^*$ tal que $|\sigma| \geq 1$, $\mathrm{SE}(ab\sigma) = \mathrm{SE}(b\sigma) = \emptyset$. □

La condición de ejecutabilidad fuerte para un σ requiere que *cada* ejecución parcial de σ (incluyendo ϵ) debe ser completada. Con esto, las fórmulas en $\mathsf{L_{Kh}}$ se interpretan en el LTS de la siguiente manera.

Definición 3.7 ($\mathsf{L_{Kh}}$ en LTSs)**.** Sean $\mathfrak{L} = \langle \mathrm{S}, \mathrm{R}, \mathrm{V} \rangle$ un LTS sobre Act y Prop, y $w \in \mathrm{D}_{\mathfrak{L}}$. La relación de satisfacibilidad $\models$ entre $\mathfrak{L}$, w y las fórmulas de $\mathsf{L_{Kh}}$ se define inductivamente de la siguiente manera:

$$\begin{array}{lll}
\mathfrak{L}, w \models p & \text{sii} & p \in \mathrm{V}(w), \\
\mathfrak{L}, w \models \neg\varphi & \text{sii} & \mathfrak{L}, w \not\models \varphi, \\
\mathfrak{L}, w \models \varphi \vee \psi & \text{sii} & \mathfrak{L}, w \models \varphi \text{ o } \mathfrak{L}, w \models \psi, \\
\mathfrak{L}, w \models \mathsf{Kh}(\psi, \varphi) & \text{sii} & \text{existe } \sigma \in \mathsf{Act}^* \text{ tal que} \\
& & \textbf{(Kh-1) } [\![\psi]\!]^{\mathfrak{L}} \subseteq \mathrm{SE}(\sigma) \text{ y } \textbf{(Kh-2) } \mathrm{R}_\sigma([\![\psi]\!]^{\mathfrak{L}}) \subseteq [\![\varphi]\!]^{\mathfrak{L}},
\end{array}$$

con $[\![\chi]\!]^{\mathfrak{L}} := \{w \in \mathrm{S} \mid \mathfrak{L}, w \models \chi\}$. Una fórmula φ es satisfacible sii $\mathfrak{L}, w \models \varphi$ para algún $\mathfrak{L}$ y $w \in \mathrm{D}_{\mathfrak{L}}$, y es válida (denotado como $\models \varphi$) sii $\mathfrak{L}, w \models \varphi$ para todo $\mathfrak{L}$ y $w \in \mathrm{D}_{\mathfrak{L}}$. Una fórmula φ se cumple (globalmente) en un LTS $\mathfrak{L}$ ($\mathfrak{L} \models \varphi$) sii $\mathfrak{L}, w \models \varphi$ para todo $w \in \mathrm{D}_{\mathfrak{L}}$. Para $\Psi \subseteq \mathsf{L_{Kh}}$, $\mathfrak{L}, w \models \Psi$ sii $\mathfrak{L}, w \models \psi$ para todo $\psi \in \Psi$; y $\Psi \models \varphi$ sii $\mathfrak{L}, w \models \Psi$ implica $\mathfrak{L}, w \models \varphi$ para todo $\mathfrak{L}$ y $w \in \mathrm{D}_{\mathfrak{L}}$. El plan σ que hace verdadero $\mathsf{Kh}(\psi, \varphi)$ en $\mathfrak{L}$ es llamado *testigo* de $\mathsf{Kh}(\psi, \varphi)$ en $\mathfrak{L}$.

Ejemplo 3.7. Se tiene que $[\![\bot]\!]^{\mathfrak{L}} = [\![\neg(p \vee \neg p)]\!]^{\mathfrak{L}} = \emptyset$ pues $\mathfrak{L}, w \models \bot$ sii $\mathfrak{L}, w \not\models (p \vee \neg p)$. Lo cual equivale a $\mathfrak{L}, w \not\models p$ y $\mathfrak{L}, w \not\models \neg p$. Esto es lo mismo que decir que $p \notin \mathrm{V}(p)$ y $p \in \mathrm{V}(p)$, una

Bloque $\mathcal{L}$:	TAUT	$\vdash \varphi$ para φ una tautología proposicional
	DISTA	$\vdash \mathsf{A}(\varphi \to \psi) \to (\mathsf{A}\varphi \to \mathsf{A}\psi)$
	TA	$\vdash \mathsf{A}\varphi \to \varphi$
	4KhA	$\vdash \mathsf{Kh}(\psi, \varphi) \to \mathsf{AKh}(\psi, \varphi)$
	5KhA	$\vdash \neg\mathsf{Kh}(\psi, \varphi) \to \mathsf{A}\neg\mathsf{Kh}(\psi, \varphi)$
Bloque $\mathcal{L}_{\text{LTS}}$:	EMP	$\vdash \mathsf{A}(\psi \to \varphi) \to \mathsf{Kh}(\psi, \varphi)$
	COMPKh	$\vdash (\mathsf{Kh}(\psi, \varphi) \wedge \mathsf{Kh}(\varphi, \chi)) \to \mathsf{Kh}(\psi, \chi)$
Reglas:	$\dfrac{\varphi \quad (\varphi \to \psi)}{\psi}$ MP	$\dfrac{\varphi}{\mathsf{A}\varphi}$ NECA

Tabla 3.1: Axiomatización $\mathcal{L}_{\mathsf{Kh}}^{\text{LTS}}$ de L_{Kh} con respecto a los LTSs.

contradicción. Por ende, no existe $w \in \mathrm{S}$ tal que $w \in [\![\bot]\!]^{\mathfrak{L}}$ y $[\![\bot]\!]^{\mathfrak{L}} = \emptyset$. Tomando el Ejemplo 3.4, $\mathfrak{L}, w_1 \models p$ pues $p \in \mathrm{V}(w_1)$. Similarmente, $\mathfrak{L}, w_2 \models r$, $\mathfrak{L}, w_4 \models r$ y $\mathfrak{L}, w_3 \models q$. Para el operador de Kh se tiene que:

- $\mathfrak{L}, w_1 \models \mathsf{Kh}(p, r)$ pues existe $\sigma = a$ tal que $[\![p]\!]^{\mathfrak{L}} = \{w_1\} \subseteq \{w_1\} = \mathrm{SE}(a)$ y $\mathrm{R}_a([\![p]\!]^{\mathfrak{L}}) = \{w_2, w_4\} \subseteq \{w_2, w_4\} = [\![r]\!]^{\mathfrak{L}}$.
- $\mathfrak{L}, w_1 \not\models \mathsf{Kh}(p, q)$ pues los únicos planes σ tales que $\{w_1\} = [\![p]\!]^{\mathfrak{L}} \subseteq \mathrm{SE}(\sigma)$ son ϵ y a. Pero $\mathrm{R}_\epsilon(w_1) = \{w_1\} \not\subseteq \{w_3\} = [\![q]\!]^{\mathfrak{L}}$ y $\mathrm{R}_a(w_1) = \{w_2, w_4\} \not\subseteq \{w_3\} = [\![q]\!]^{\mathfrak{L}}$. Luego, no existe $\sigma \in \mathsf{Act}^*$ tal que $[\![p]\!]^{\mathfrak{L}} \subseteq \mathrm{SE}(\sigma)$ y $\mathrm{R}_\sigma([\![p]\!]^{\mathfrak{L}}) \subseteq [\![q]\!]^{\mathfrak{L}}$.

$\square$

De esta manera, $\mathsf{Kh}(\psi, \varphi)$ es verdadero en un estado w cuando existe un plan tal que, al ser ejecutado en cualquier estado donde se cumpla ψ, siempre completará cada ejecución parcial (condición **(Kh-1)**), terminando sin fallas en estados que satisfagan φ (condición **(Kh-2)**). Como w no influye en la semántica de Kh, el operador de *saber cómo* actúa *globalmente*. Por ende, $[\![\mathsf{Kh}(\psi, \varphi)]\!]^{\mathfrak{L}}$ es siempre $\mathrm{D}_{\mathfrak{L}}$ ó $\emptyset$, y además $\mathfrak{L}, w \models \mathsf{Kh}(\psi, \varphi)$ sii $\mathfrak{L} \models \mathsf{Kh}(\psi, \varphi)$.

3.2. Axiomatización

En esta sección se presenta un sistema axiomático para la lógica L_{Kh}. Como se puede observar en la Tabla 3.1, la misma depende fuertemente de la modalidad universal A [51]. Esta es una modalidad unaria que representa las propiedades que se cumplen en todos los estados del modelo. Es decir, las fórmulas de la forma $\mathsf{A}\varphi$ se leen como *"φ se cumple en todos los estados"*. Como se demostró en [116], A es definible en términos de Kh como $\mathsf{A}\varphi := \mathsf{Kh}(\neg\varphi, \bot)$. La proposición que sigue justifica este hecho, y la prueba se sostiene en que Act^* es no vacío, ya que $\epsilon \in \mathsf{Act}^*$. Por lo tanto, puede tomárselo como testigo en la cláusula existencial en la semántica de Kh.

Proposición 3.1 ([116]). *Sean $\mathfrak{L}$ un* LTS *y $w \in \mathrm{D}_{\mathfrak{L}}$. Entonces,*

$$\mathfrak{L}, w \models \mathsf{Kh}(\neg\varphi, \bot) \qquad \textit{sii} \qquad [\![\varphi]\!]^{\mathfrak{L}} = \mathrm{D}_{\mathfrak{L}}.$$

Ejemplo 3.8. Tomando el modelo del Ejemplo 3.4, como $[\![(p \vee q \vee r)]\!]^{\mathfrak{L}} = \mathrm{S}$, tomando $\sigma = \epsilon$, $[\![\neg(p \vee q \vee r)]\!]^{\mathfrak{L}} = ([\![(p \vee q \vee r)]\!]^{\mathfrak{L}})^c = \emptyset \subseteq \mathrm{S} = \mathrm{SE}(\epsilon)$ y $\mathrm{R}_\epsilon([\![\neg(p \vee q \vee r)]\!]^{\mathfrak{L}}) = [\![\neg(p \vee q \vee r)]\!]^{\mathfrak{L}} = \emptyset \subseteq \emptyset = [\![\bot]\!]^{\mathfrak{L}}$ por el Ejemplo 3.7. Por lo tanto, $\mathfrak{L}, w_1 \models \mathsf{A}(p \vee q \vee r)$. $\square$

El sistema axiomático $\mathcal{L}_{\mathsf{Kh}}^{\text{LTS}}$ (Tabla 3.1) muestra la relación entre la modalidad global universal A y el operador de saber cómo Kh. El primer bloque y las reglas MP y NECA, son esencialmente el sistema modal estándar para A [51], estableciendo además que Kh es global. En el caso del segundo bloque, el axioma EMP establece que, si $\psi \to \varphi$ es globalmente verdadera, entonces dado ψ el agente sabe cómo hacer φ verdadera. Esto es debido a que el plan vacío ϵ siempre está disponible. Con ello, la información global óntica se vuelve conocimiento. Mientras tanto, el axioma COMP describe que Kh es composicional: si dado ψ el agente sabe cómo hacer φ verdadera, y dado φ el agente sabe cómo hacer χ verdadera, entonces dado ψ sabe cómo hacer χ verdadera.

Ejemplo 3.9. Usando el LTS $\mathfrak{L}$ del Ejemplo 3.4, se cumple $\mathfrak{L} \models \mathsf{A}((\neg r \wedge \neg q) \to p)$. Aplicando la Definición 3.7, para todo $w \in \mathrm{S}$, $\mathfrak{L}, w \models (\neg r \wedge \neg q)$ implica $\mathfrak{L}, w \models p$. Esto es equivalente a que $[\![(\neg r \wedge \neg q)]\!]^{\mathfrak{L}} \subseteq [\![p]\!]^{\mathfrak{L}}$. Utilizando la definición de R_ϵ, como $\mathrm{R}_\epsilon([\![(\neg r \wedge \neg q)]\!]^{\mathfrak{L}}) = [\![(\neg r \wedge \neg q)]\!]^{\mathfrak{L}} \subseteq [\![p]\!]^{\mathfrak{L}}$ y $\mathrm{SE}(\epsilon) = \mathrm{S}$, existe un plan testigo $\sigma = \epsilon$ tal que $[\![(\neg r \wedge \neg q)]\!]^{\mathfrak{L}} \subseteq \mathrm{SE}(\sigma)$ y $\mathrm{R}_\sigma([\![(\neg r \wedge \neg q)]\!]^{\mathfrak{L}}) \subseteq [\![p]\!]^{\mathfrak{L}}$. Por ende, $\mathfrak{L} \models \mathsf{Kh}((\neg r \wedge \neg q), p)$ y con esto $\mathfrak{L} \models \mathsf{A}((\neg r \wedge \neg q) \to p) \to \mathsf{Kh}((\neg r \wedge \neg q), p)$.

Dado que $\mathfrak{L} \models \mathsf{Kh}(p, r)$ con $\sigma' = a$ como testigo, componiendo ambos planes se tiene el testigo $\sigma'' = \sigma\sigma' = \epsilon a = a$ tal que $[\![(\neg r \wedge \neg q)]\!]^{\mathfrak{L}} \subseteq [\![p]\!]^{\mathfrak{L}} = \{w_1\} \subseteq \mathrm{SE}(a)$ y $\mathrm{R}_a([\![(\neg r \wedge \neg q)]\!]^{\mathfrak{L}}) \subseteq \mathrm{R}_a([\![p]\!]^{\mathfrak{L}}) = \{w_2, w_4\} \subseteq [\![r]\!]^{\mathfrak{L}}$. Por lo tanto, $\mathfrak{L} \models \mathsf{Kh}((\neg r \wedge \neg q), r)$. $\square$

Esta axiomatización además es correcta y fuertemente completa [116]. Lo que quiere decir que ***(1)*** todos los teoremas deducibles del sistema son válidos en todos los LTSs, y ***(2)*** todas las fórmulas válidas en todos los LTSs son derivables de estos axiomas. Estas características son importantes ya que vincula la noción de consecuencia sintáctica ($\vdash$) con la noción de consecuencia semántica ($\models$), haciéndolas coincidir de manera tal que todo teorema es una tautología y toda tautología es un teorema.

Teorema 3.1 ([116])**.** *El sistema axiomático $\mathcal{L}_{\mathsf{Kh}}^{\mathrm{LTS}}$ (Tabla 3.1) es correcto y fuertemente completo para L_{Kh} con respecto a la clase de todos los* LTS*s.*

A lo largo de este trabajo se verán extensiones o modificaciones de esta axiomatización para describir diferentes lógicas inspiradas en L_{Kh} (ver Capítulos 5, 9 y 10). Por ejemplo, pese a que L_{Kh} posee cierto consenso como una lógica adecuada para modelar la noción de saber cómo, los axiomas en el segundo bloque pueden ser objeto de cuestionamientos. En primer lugar, se podría argumentar que, en contraste con el axioma EMP, no todas las verdades globales sobre qué es alcanzable necesitan ser consideradas como conocimiento del agente. Por ejemplo, no todas las verdades del universo pueden transformarse en conocimiento para un niño. Es más, dichas verdades pueden no influir en nada en dicho conocimiento. En segundo lugar, el axioma COMPKh también implica un cierto nivel de omnisciencia: es posible que un agente sepa cómo hacer φ verdadera dado ψ, y sepa cómo hacer χ verdadera dado φ, pero que todavía no haya logrado componer los dos testigos para asegurar χ dado ψ. Una vez más, un niño puede saber cómo ir de un punto A a un punto B y de un punto B a un punto C, pero no tener la suficiente introspección para utilizar ambas rutas y saber cómo ir de A a C. Estas dos propiedades serán discutidas nuevamente al introducir una nueva semántica en el Capítulo 4.

3.3. Bisimulaciones

A la hora de comparar el comportamiento entre dos sistemas de transiciones etiquetadas distintos, una de las herramientas disponibles es la noción de bisimulación [95,96]. Dos LTSs son bisimilares (existe una bisimulación entre estos) si ambos sistemas se comportan de la misma manera. Es decir que, de acuerdo a la noción usual, si un sistema realiza una acción cualquiera, el otro podrá emular la misma y viceversa. Lo cual implica que no existe una serie de acciones las cuales un sistema puede ejecutarlas y el otro no. De esta forma las bisimulaciones son una herramienta relevante para entender el poder expresivo de una lógica modal.

En términos lógicos, la noción de bisimulación depende del lenguaje subyacente. Dos LTSs son bisimilares cuando el comportamiento de ambos no puede ser distinguido por el lenguaje. Cada vez que se busca definir una bisimulación para una lógica específica, se debe tener cuidado a la hora de formular las condiciones: deben ser lo suficientemente fuertes para garantizar que la lógica no pueda distinguir dos modelos bisimilares, pero lo suficientemente débiles para que valgan en modelos que no pueden ser distinguibles por la lógica, es decir que no exista una fórmula que se cumpla en un modelo y no en el otro. De esta manera, podemos asegurar que la noción de la bisimulación es suficiente para caracterizar la expresividad de la lógica.

En el caso de la lógica L_{Kh}, la noción de bisimulación de, por ejemplo, la lógica básica multimodal (con operadores $[a]$, $a \in \mathsf{Act}$) no es conveniente. Esto es debido a que el operador Kh no habla de los planes explícitamente. Más aún, la modalidad universal ni siquiera es definible en dicha lógica. Por lo tanto, se deben considerar otras alternativas que contemplen estos dos aspectos. Esto sucede en [43,44], donde se hablan ya no de acciones básicas o planes específicos, y estados bisimilares, sino de la *existencia de planes fuertemente ejecutables* y de *conjuntos de estados* bisimilares. Considerando esto, en esta sección se establecerán los conceptos principales de bisimulaciones para LTSs.

Definición 3.8. Sea $\mathfrak{L} = \langle \mathrm{S}, \mathrm{R}, \mathrm{V} \rangle$ un LTS sobre Prop y Act. Sean $\sigma \in \mathsf{Act}^*$ un plan y $U, T \subseteq \mathrm{S}$ conjuntos de estados.

- $U \overset{\sigma}{\Rightarrow} T$ sii $U \subseteq \mathrm{SE}(\sigma)$ y $\mathrm{R}_\sigma(U) = T$.
- $U \Rightarrow T$ sii existe $\sigma \in \mathsf{Act}^*$ tal que $U \overset{\sigma}{\Rightarrow} T$.

Además, $U \subseteq \mathrm{S}$ es $\mathsf{L_{Kh}}$-definible (respectivamente proposicionalmente definible) en $\mathfrak{L}$ si y sólo si existe una fórmula en $\mathsf{L_{Kh}}$ (fórmula proposicional) φ tal que $U = [\![\varphi]\!]^{\mathfrak{L}}$.

Ejemplo 3.10. Tomando el LTS del Ejemplo 3.4, se puede observar que $\{w_1\} \overset{a}{\Rightarrow} \{w_2, w_4\}$ y $\{w_2\} \overset{b}{\Rightarrow} \{w_3\}$, pero no ocurre que $\{w_1\} \overset{ab}{\Longrightarrow} \{w_3\}$ y $\{w_2, w_4\} \overset{b}{\Rightarrow} \{w_3\}$. Además de eso, $\{w_2, w_4\}$ es proposicionalmente definible por $\varphi = r$. Sin embargo no es el caso de $\{w_2\}$ ya que w_2 y w_4 comparten todos los valores de verdad sobre todas las variables proposicionales. $\square$

Dada la semántica global del operador Kh, se puede ver que $\mathsf{L_{Kh}}$-definibilidad implica definibilidad proposicional.

Proposición 3.2 ([43, 44])**.** *Sea $\mathfrak{L}$ un* LTS*. Para todo $U \subseteq \mathrm{D}_{\mathfrak{L}}$, si U es $\mathsf{L_{Kh_i}}$-definible, entonces es proposicionalmente definible.*

Una vez introducidas las nociones anteriores, es posible definir la noción de bisimulación para la lógica $\mathsf{L_{Kh}}$.

Definición 3.9 ($\mathsf{L_{Kh}}$-bisimulación)**.** Sean $\mathfrak{L}$ y $\mathfrak{L}'$ dos LTSs, tal que $\mathrm{D}_{\mathfrak{L}} = \mathrm{S}$ y $\mathrm{D}_{\mathfrak{L}'} = \mathrm{S}'$. Sea $Z \subseteq \mathrm{S} \times \mathrm{S}'$.

- Para todo $u \in \mathrm{S}$ y $U \subseteq \mathrm{S}$, se definen
$$Z(u) := \{u' \in \mathrm{S}' \mid uZu'\}, \qquad Z(U) := \textstyle\bigcup_{u \in U} Z(u).$$
- Para todo $u' \in \mathrm{S}'$ y $U' \subseteq \mathrm{S}'$, se definen
$$Z^{-1}(u') := \{u \in \mathrm{S} \mid uZu'\}; \qquad Z^{-1}(U') := \textstyle\bigcup_{u' \in U'} Z^{-1}(u').$$

Una relación no vacía $Z \subseteq \mathrm{S} \times \mathrm{S}'$ es una $\mathsf{L_{Kh}}$-bisimulación entre $\mathfrak{L}$ y $\mathfrak{L}'$ si y sólo si wZw' implica lo siguiente.

- **Atom**: $\mathrm{V}(w) = \mathrm{V}'(w')$.
- Kh-**Zig**: para todo conjunto *proposicionalmente* definible $U \subseteq \mathrm{S}$, si $U \Rightarrow T$ para algún $T \subseteq \mathrm{S}$, entonces existe $T' \subseteq \mathrm{S}'$ tal que

  (B1) $Z(U) \Rightarrow T'$, ***(B2)*** $T' \subseteq Z(T)$.

- Kh_i-**Zag**: para todo conjunto *proposicionalmente* definible $U' \subseteq \mathrm{S}'$, si $U' \Rightarrow T'$ para algún $T' \subseteq \mathrm{S}'$, entonces existe $T \subseteq \mathrm{S}$ tal que

  (B1) $Z^{-1}(U') \Rightarrow T$, ***(B2)*** $T \subseteq Z^{-1}(T')$.

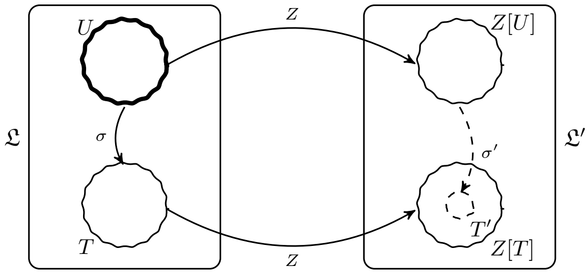
- A-**Zig**: para todo $u \in \mathrm{S}$ existe un $u' \in \mathrm{S}'$ tal que uZu'.
- A-**Zag**: para todo $u' \in \mathrm{S}'$ existe un $u \in \mathrm{S}$ tal que uZu'.

Se escribe $\mathfrak{L}, w \leftrightarroweq_{\mathsf{L_{Kh}}} \mathfrak{L}', w'$ cuando existe una $\mathsf{L_{Kh}}$-bisimulación Z entre $\mathfrak{L}$ y $\mathfrak{L}'$ tal que wZw'.

El siguiente diagrama muestra las condiciones que Kh-**Zig** impone. Sea $U \subseteq \mathrm{S}$ un subconjunto proposicionalmente definible, si $U \Rightarrow T$ (existe $\sigma \in \mathsf{Act}^*$ tal que $U \subseteq \mathrm{SE}(\sigma)$ y $\mathrm{R}_\sigma(U) = T$), entonces existe un $T' \subseteq \mathrm{S}'$ tal que

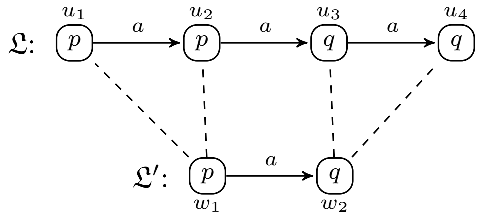
(1) $Z(U) \Rightarrow T'$, es decir, para algún $\sigma' \in \mathsf{Act}^*$, T' tiene exactamente todos los estados que son alcanzables por algún estado en $Z(U)$ vía σ' ($\mathrm{R}_{\sigma'}(Z(U)) = T'$), y σ es fuertemente ejecutable para todos los estados en $Z(U)$ ($Z(U) \subseteq \mathrm{SE}(\sigma')$); y

(2) cada estado en T' es la imagen de algún estado en T vía Z ($T' \subseteq Z(T)$).



Las condiciones de Kh-**Zag** funcionan en la otra dirección: para cualquier conjunto proposicionalmente definible $U' \subseteq \mathrm{S}'$, que cumpla $U' \Rightarrow T'$ para algún $T' \subseteq \mathrm{S}'$, entonces existe $T \subseteq \mathrm{S}$ tal que *(1)* $Z^{-1}(U') \Rightarrow T$ y *(2)* $T \subseteq Z^{-1}(T')$. El diagrama es análogo al anterior.

Ejemplo 3.11. Sean $\mathfrak{L} = \langle \mathrm{S}, \mathrm{R}, \mathrm{V} \rangle$ y $\mathfrak{L}' = \langle \mathrm{S}', \mathrm{R}', \mathrm{V}' \rangle$ los siguientes LTSs.



Tomando $Z = \{(u_1, w_1), (u_2, w_1), (u_3, w_2), (u_4, w_2)\}$, representado en el gráfico con las líneas punteadas, se puede observar que Z cumple **Atom**, A-**Zig** y A-**Zag**. Más aún, Kh-**Zig** también. Todo subconjunto $U \subseteq \mathrm{S}$ proposicionalmente definible satisface las condiciones como se muestra en la siguiente tabla.

U	Def. como	T	$U \Rightarrow T$ por	$Z(U)$	$Z(U) \Rightarrow T'$ por	T'	$Z(T)$
$\emptyset$	$\bot$	$\emptyset$	ϵ	$\emptyset$	ϵ	$\emptyset$	$\emptyset$
$\{u_1, u_2\}$	p	$\{u_1, u_2\}$	ϵ	$\{w_1\}$	ϵ	$\{w_1\}$	$\{w_1\}$
		$\{u_2, u_3\}$	a	$\{w_1\}$	a	$\{w_2\}$	S'
		$\{u_3, u_4\}$	aa	$\{w_1\}$	a	$\{w_2\}$	$\{w_2\}$
$\{u_3, u_4\}$	q	$\{u_3, u_4\}$	ϵ	$\{w_2\}$	ϵ	$\{w_2\}$	$\{w_2\}$
S	$\top$	S	ϵ	S'	ϵ	S'	S'

En la misma, se listan exhaustivamente todos los conjuntos U, T y T' en el modelo y sus correspondientes imágenes $Z(U)$ y $Z(T)$. La segunda columna especifica qué fórmula proposicional define el conjunto U mientras que la cuarta y sexta columnas enumeran los testigos para cada relación $U \Rightarrow T$ y $Z(U) \Rightarrow T'$ respectivamente. Notar que en cada fila $T' \subseteq Z(T)$. Con un razonamiento similar se obtiene Kh-**Zag**. Con esto, Z es una $\mathsf{L_{Kh}}$-bisimulación entre $\mathfrak{L}$ y $\mathfrak{L}'$. Además, $\mathfrak{L}, u_1 \,\underline{\leftrightarrow}_{\mathsf{L_{Kh}}}\, \mathfrak{L}', w_1$ y $\mathfrak{L}, u_4 \,\underline{\leftrightarrow}_{\mathsf{L_{Kh}}}\, \mathfrak{L}', w_2$ dado que $u_1 Z w_1$ y $u_4 Z w_2$ respectivamente. $\square$

En lógica modal, la definición bisimilaridad está conectada con la de equivalencia de fórmulas. Esto es debido a que, dados dos modelos que se comportan de manera similar desde la perspectiva de una lógica modal determinada, deben cumplir las mismas propiedades desde esa lógica.

Definición 3.10 ($\mathsf{L_{Kh}}$-equivalencia)**.** Sean $\mathfrak{L}$ y $\mathfrak{L}'$ LTSs, $w \in \mathrm{D}_{\mathfrak{L}}$ y $w' \in \mathrm{D}_{\mathfrak{L}'}$ estados. $\mathfrak{L}, w$ y $\mathfrak{L}', w'$ son $\mathsf{L_{Kh}}$-*equivalentes* ($\mathfrak{L}, w \leftrightsquigarrow_{\mathsf{L_{Kh}}} \mathfrak{L}', w'$) si y sólo si, para cada $\varphi \in \mathsf{L_{Kh_i}}$,

$$\mathfrak{L}, w \models \varphi \quad \text{sii} \quad \mathfrak{L}', w' \models \varphi.$$

Con esto se puede establecer la correspondencia entre $\underline{\leftrightarrow}_{\mathsf{L_{Kh}}}$ y $\leftrightsquigarrow_{\mathsf{L_{Kh}}}$.

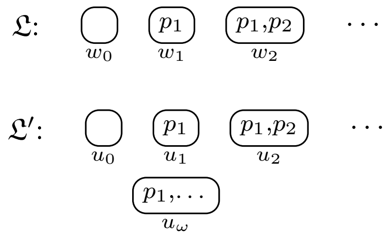
Teorema 3.2 (Invarianza bajo bisimulaciones [43,44])**.** *Sean $\mathfrak{L}$ y $\mathfrak{L}'$ LTSs, $w \in \mathrm{D}_{\mathfrak{L}}$ y $w' \in \mathrm{D}_{\mathfrak{L}'}$ estados. Entonces,*

$$\mathfrak{L}, w \,\underline{\leftrightarrow}_{\mathsf{L_{Kh}}}\, \mathfrak{L}', w' \quad \textit{implica} \quad \mathfrak{L}, w \leftrightsquigarrow_{\mathsf{L_{Kh}}} \mathfrak{L}', w'.$$

Ejemplo 3.12. Tomando los modelos del Ejemplo 3.11, como $\mathfrak{L}, u_1 \underline{\leftrightarrow}_{\mathsf{L_{Kh}}} \mathfrak{L}', w_1$ y $\mathfrak{L}, u_4 \underline{\leftrightarrow}_{\mathsf{L_{Kh}}} \mathfrak{L}', w_2$, entonces $\mathfrak{L}, u_1 \leftrightsquigarrow_{\mathsf{L_{Kh}}} \mathfrak{L}', w_1$ y $\mathfrak{L}, u_4 \leftrightsquigarrow_{\mathsf{L_{Kh}}} \mathfrak{L}', w_2$. □

Desafortunadamente, la vuelta no siempre se cumple para modelos arbitrarios. Esto se debe en parte a que la modalidad universal A es definible en esta lógica. Para demostrar este punto, se tiene el contraejemplo provisto en [44, Sección 2]. Lo siguiente es una leve modificación del mismo.

Ejemplo 3.13. Sean $\mathfrak{L} = \langle \mathrm{S}, \mathrm{R}, \mathrm{V} \rangle$ y $\mathfrak{L}' = \langle \mathrm{S}', \mathrm{R}', \mathrm{V}' \rangle$, $\mathsf{Prop} = \{p_i \mid i \in \mathbb{N}\}$ tal que $\mathfrak{L}$ y $\mathfrak{L}$ tienen una cantidad infinita numerable de estados $\mathrm{S} = \{w_0, w_1, w_2, \ldots\}$ y $\mathrm{S}' = \{u_\omega, u_0, u_1, u_2, \ldots\}$. Supongamos que $\mathrm{V}(w_n) = \mathrm{V}'(u_n)$, $p_k \in \mathrm{V}(w_n)$ si y sólo si $n \geq k$ y $\mathrm{V}'(u_\omega) = \mathsf{Prop}$. Además, $\mathrm{R} = \mathrm{R}' = \emptyset$.



Notar que $\mathfrak{L}, w_0 \leftrightsquigarrow_{\mathsf{L_{Kh}}} \mathfrak{L}', u_0$. Por definición, ambos estados coinciden en las variables proposicionales. Además, cualquier fórmula de la forma $\mathsf{Kh}_i(\psi, \varphi)$ que se cumpla en $\mathfrak{L}$ (con ϵ como único testigo posible), se cumple en $\mathfrak{L}'$ con el mismo testigo y viceversa. Esto es posible ya que las fórmulas ψ y φ son finitas.

Sin embargo, no existe una bisimulación tal que $\mathfrak{L}, w_0 \underline{\leftrightarrow}_{\mathsf{L_{Kh}}} \mathfrak{L}', u_0$. De existir dicha bisimulación Z, u_ω no se podría emparejar con ningún estado de S dado que ninguno coincide proposicionalmente con este. Pero esto violaría la condición de A-**Zag**, pues no existiría $w_n \in \mathrm{S}$ tal que $w_n Z u_\omega$. Por lo tanto, $\mathfrak{L}, w_0 \underline{\leftrightarrow}_{\mathsf{L_{Kh}}} \mathfrak{L}', u_0$ no sucede. □

Para que se cumpla la vuelta, usualmente se tiene que restringir a una clase particular de modelos conocidos como clases de Hennessy-Milner. En el caso de las lógicas modales, esta clase son de los modelos con imagen finita [26]. Es decir, modelos tales que por cada estado y por cada relación básica, se tiene una cantidad finita de sucesores. Sin embargo, para lógicas en las que la modalidad universal A es definible, como sucede con $\mathsf{L_{Kh}}$, es necesario fortalecer las condiciones ya que cada estado puede alcanzar a cualquier otro. Por ende, una buena candidata resulta ser la clase de modelos con dominio finito. Con esto se tiene la conversa del teorema anterior.

Teorema 3.3 ($\mathsf{L_{Kh}}$-equivalencia implica $\mathsf{L_{Kh}}$-bisimilaridad [43, 44])**.** *Sean $\mathfrak{L}$ y $\mathfrak{L}'$ LTSs finitos, $w \in \mathrm{D}_{\mathfrak{L}}$ y $w' \in \mathrm{D}_{\mathfrak{L}'}$ estados. Entonces,*

$$\mathfrak{L}, w \leftrightsquigarrow \mathfrak{L}', w' \quad \textit{implica} \quad \mathfrak{L}, w \underline{\leftrightarrow} \mathfrak{L}', w'.$$

Generalmente el resultado de estos teoremas facilita la tarea de comparar la expresividad entre dos lógicas modales. Si se toman dos modelos indistinguibles bajo una determinada lógica y se puede demostrar que una fórmula de otra lógica distingue un modelo del otro, entonces la primera lógica no es igual o más expresiva que la segunda. Si esto también sucede al revés, se tiene que ambas son incomparables.

3.4. Complejidad

A lo largo de esta tesis nos enfocaremos en describir la complejidad computacional de dos problemas de decisión importantes: el problema model checking (o verificación de modelos) y el problema de satisfacibilidad. El primero se formula como "dado un modelo y una fórmula de la lógica, determinar si dicha fórmula se cumple o no en el modelo" mientras que el segundo es de la forma "dada una fórmula de la lógica, determinar si existe un modelo que la satisfaga".

La clase de complejidad computacional a la que pertenecen ambos problemas varía según la lógica que se esté estudiando. Por ejemplo, para la lógica proposicional, model checking está en P mientras que satisfacibilidad está en NP [12, 35]. Para lógica modal básica, que agrega expresividad hablando de mundos posibles, el primero está en P y el segundo es PSpace-completo.

Por otra parte, para la lógica de primer orden, que contiene a las anteriores, model checking es PSpace-completo pero el problema de satisfacibilidad es indecidible [25]. Con esto, se puede inferir una correlación: cuanto más expresiva sea una lógica, sus problemas de model checking y satisfacibilidad pertenecerán a una clase de complejidad más alta.

En el caso de $\mathsf{L_{Kh}}$, ambos problemas son decidibles. Específicamente, en [36], se demuestra que model checking es PSpace-completo mediante la construcción de autómatas finitos determinísticos y lenguajes regulares sobre el conjunto de planes testigo de una fórmula $\mathsf{Kh}(\varphi_1, \varphi_2)$ en un LTS dado.

Teorema 3.4 ([36]). *El problema de model checking para* $\mathsf{L_{Kh}}$ *es* PSpace-*completo.*

Por otro lado, para el problema de satisfacibilidad, fue demostrado que tiene $\mathsf{NP^{NP}}$ como cota superior [6], una clase de complejidad entre NP y PSpace que utiliza un oráculo para determinar si una serie de fórmulas proposicionales son satisfacibles o no [12].

Teorema 3.5 ([6]). *El problema de satisfacibilidad para* $\mathsf{L_{Kh}}$ *está en* $\mathsf{NP^{NP}}$.

Vale la pena destacar que usualmente satisfacibilidad se encuentra en una clase de complejidad igual o más alta que la del problema de model checking. Sin embargo, no es el caso para esta lógica. Esto se debe a que la satisfacibilidad de una fórmula en $\mathsf{L_{Kh}}$ depende exclusivamente de la satisfacibilidad de un conjunto subyacente de fórmulas proposicionales. Mientras tanto, para model checking es necesario almacenar en espacio polinomial parte de los autómatas construidos a partir del LTS y la fórmula dadas. Estos problemas de decidibilidad se retomarán a lo largo de esta tesis. Especialmente, en los Capítulos 7, 9 y 10.

3.5. Otros operadores de saber cómo

Además de la lógica $\mathsf{L_{Kh}}$ discutida a lo largo de este capítulo, hay otras variaciones de la modalidad de saber cómo. Así como es el caso de [72, 76] que propone un operador ternario $\mathsf{Khm}(\psi, \chi, \varphi)$ el cual además de las condiciones originales, se pide que se cumplan condiciones intermedias. Es decir que todos los estados por los que pase el plan testigo cumplan determinadas propiedades. La semántica de esta modalidad, también interpretada sobre LTSs, es la siguiente:

$\mathfrak{L}, w \models \mathsf{Khm}(\psi, \chi, \varphi)$ sii existe $\sigma \in \mathsf{Act}^*$ tal que **(Khm-1)** $[\![\psi]\!]^{\mathfrak{L}} \subseteq \mathrm{SE}(\sigma)$, **(Khm-2)** $\mathrm{R}_\sigma([\![\psi]\!]^{\mathfrak{L}}) \subseteq [\![\varphi]\!]^{\mathfrak{L}}$ y **(Khm-3)** $\mathrm{R}_{\sigma_k}([\![\psi]\!]^{\mathfrak{L}}) \subseteq [\![\chi]\!]^{\mathfrak{L}}$ para todo $k \in [1 \ldots |\sigma| - 1]$.

Otra propuesta es el caso de [72, 73] con la definición de $\mathsf{Khw}(\psi, \varphi)$, en la cual se debilita la condición de ejecutabilidad fuerte. En vez de pedir que todas las ejecuciones parciales se completen y los estados finales cumplan φ, se pide al menos que en los estados donde las ejecuciones terminen (sean completas o no) se cumpla φ. La semántica se define de la siguiente manera:

$$\mathfrak{L}, w \models \mathsf{Khw}(\psi, \varphi) \text{ sii existe } \sigma \in \mathsf{Act}^* \text{ tal que para cada } w \in [\![\psi]\!]^{\mathfrak{L}}, \mathtt{TermiSs}(w, \sigma) \subseteq [\![\varphi]\!]^{\mathfrak{L}}$$

con $\mathtt{TermiSs}(w, \sigma)$ siendo el conjunto de todos los estados terminales en donde las ejecuciones de σ terminan parcial o completamente. Es decir, $u \in \mathtt{TermiSs}(w, \sigma)$ si y sólo si $u \in \mathrm{R}_\sigma(w)$ ó $u \in \mathrm{R}_{\sigma_k}(w)$ y $\mathrm{R}_{\sigma[k]}(u) = \emptyset$ para algún $k \in [0 \ldots |\sigma| - 1]$. Finalmente, en [115], se tiene una modalidad $\mathsf{Khs}(\psi, \varphi)$ en la cual se da otro debilitamiento de la condición de ejecutabilidad fuerte. En este caso se evalúa sobre los estados que llegan después de ejecutar σ desde w.

$$\mathfrak{L}, w \models \mathsf{Khs}(\psi, \varphi) \text{ sii existe } \sigma \in \mathsf{Act}^* \text{ tal que para cada } w \in [\![\psi]\!]^{\mathfrak{L}}, \mathtt{ArrSta}(w, \sigma) \subseteq [\![\varphi]\!]^{\mathfrak{L}}.$$

Es decir, $u \in \mathtt{ArrSta}(w, \sigma)$ si y sólo si $u \in \mathrm{R}_\sigma(w)$ ó $u = w$ si $\mathrm{R}_\sigma(w) = \emptyset$. Para todos estas variantes, existen resultados de bisimulaciones y de expresividad entre estas lógicas en [44, 115]. En el apartado de los problemas de decisión de satisfabilidad, al momento de escribir este trabajo, sólo se han probado que son decidibles [72].

Cabe destacar que todas estas modalidades mencionadas son globales. O se cumplen en todos los estados o en ninguno. Sin embargo, existen variantes de saber cómo en las cuales el operador ha sido definido de manera local como en [42] donde incluso se lo combina con una modalidad de saber qué. En dicho trabajo se provee una axiomatización y resultados de decidibilidad para el problema de satisfabilidad. Estos últimos se refinan en [74, 75], donde el problema de satisfabilidad es PSpace-completo. Más aún, se da una generalización en [77] donde se prueban propiedades similares. Otros trabajos relacionados con nociones de saber cómo se encuentran en [85–87] vinculado a las ideas de “estrategias de coalición”.

Capítulo 4

Lógica de saber cómo basada en incertidumbre

> Existe más de un sueño en cada vida. Tal vez existan tantos como años. Para mí el último es: "Saber". Creo que es el más largo y el menos realizable.
>
> Maeterlinck

En el capítulo anterior se introdujo una lógica basada en Sistemas de Transiciones Etiquetadas que provee una representación comprensible de las habilidades de un agente, junto con una interpretación sobre el conocimiento del mismo acerca de su capacidad de lograr un objetivo. El agente sabe como lograr φ dado ψ solamente cuando existe un plan tal que, al ser ejecutado en un estado donde vale ψ, siempre completará cada ejecución parcial de tal plan, llegando sólo a estados que satisfacen φ. Sin embargo, dicha representación supone que el agente posee ciertas características ideales, como ya ha sido discutido al introducir el sistema axiomático adecuado para la lógica.

Supongamos que tenemos un agente que carece de una determinada habilidad. En L_{Kh}, esta situación puede darse únicamente cuando el entorno (el modelo LTS) no tiene la secuencia de acciones requerida. Sin embargo, hay situaciones en las que un plan adecuado sí existe en el entorno, pero el agente sigue careciendo de dicha habilidad por otras razones. Puede suceder que el agente no es *capaz de distinguir* un plan adecuado de otro en el sentido de que no es consciente de que ambos planes dan diferentes resultados.

Consideremos un agente que está haciendo una torta. El agente posee la habilidad de realizar cuatro métodos de mezcla diferentes (método de batido, de la crema, fusión y frotamiento), e incluso reconocerlos como diferentes acciones. Sin embargo, puede no ser capaz de distinguirlos entre sí o reconocer que pueden producir diferentes resultados. Puede decirse en este caso que el agente no sabe cómo hacer una torta: a veces puede tener un buen resultado (al usar el método correcto) y otras no. En este caso, el agente no puede distinguir entre *acciones básicas* (los métodos de mezcla).

Aún así, se puede pensar una forma más general de indistinguibilidad que involucre *planes*, y no sólo acciones atómicas. Consideremos el agente anterior. Podemos decir que sabe la diferencia entre "poner leche" y "poner harina" a la mezcla, pero quizás no es consciente que el *orden* de estas acciones influyen en el resultado final. El problema en este caso no es que no pueda distinguir entre acciones básicas, sino que no puede distinguir dos planes, ya que considera irrelevante el orden de las acciones. Más aún, el agente puede no saber que, si bien es necesario abrir el horno para verificar si el proceso de horneado está completo, hacerlo con demasiada frecuencia puede afectar el resultado. En este caso, el agente no es capaz de distinguir entre el efecto de ejecutar una acción una vez y hacerlo varias veces. Con esto, planes de *diferente longitud* pueden ser considerados equivalentes según la perspectiva de dicho agente, cuando los mismos producen diferentes efectos.

Estos ejemplos indican que se puede crear una representación más general sobre las habilidades del agente tomando en cuenta no sólo los planes que están disponibles (en el modelo LTS) sino también la capacidad del agente de distinguir planes entre sí. Gracias a esta (in)capacidad de distinguir planes, se puede definir una representación con múltiples agentes en la cual estos compartan el mismo *entorno* pero que cada uno tenga su propia *percepción* acerca de los efectos de cada acción o secuencia de acciones. Esto último quiere decir que es posible que para el agente hayan planes tan ajenos o demasiado complejos, que quizás no sea siquiera *consciente* de ellos. En otras palabras, dichos planes están fuera del alcance del agente dado que no puede considerarlos. Esto nos permite modelar agentes humanos o computacionales menos idealizados, más 'reales' y con recursos limitados. En resumen, $\mathsf{L_{Kh}}$ no contempla la posibilidad de que el agente ***(1)*** no sea capaz de distinguir entre ciertos planes entre sí, o ***(2)*** no sea consciente de la existencia de todos los planes.

En el resto del capítulo, se definirá una lógica de saber cómo multiagente, introducida en [8, 10], que considere estas situaciones. Se reintroducirán los modelos LTS con una componente que represente la indistinguibilidad entre planes para cada uno de los agentes. Dicha lógica se denomina *basada en incertidumbre*, dado que representa la incerteza del agente a la hora de discernir ciertos planes. Esto permitirá definir escenarios con múltiples agentes los cuales comparten el mismo LTS, que representa lo que puede hacer un agente, pero con diferentes percepciones sobre sus habilidades, lo que es consciente de.

4.1. Sintaxis y semántica

En esta sección se introduce la sintaxis y la semántica de la lógica de saber cómo basada en incertidumbre, denotada como $\mathsf{L_{Kh_{\mathit{i}}}}$. El primer cambio con respecto a $\mathsf{L_{Kh}}$ es que esta lógica se consideran múltiples agentes. A partir de ahora, se considera un conjunto numerable no vacío de símbolos proposicionales Prop, un conjunto finito no vacío de agentes Agt y un conjunto infinito numerable de acciones básicas Act. A continuación se define la sintaxis de la lógica $\mathsf{L_{Kh_{\mathit{i}}}}$.

Definición 4.1. Las fórmulas de $\mathsf{L_{Kh_{\mathit{i}}}}$ son dadas por la gramática

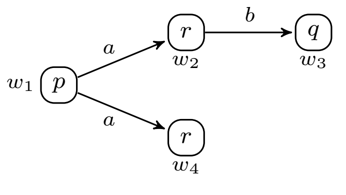
$$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid \mathsf{Kh}_i(\varphi, \varphi),$$

con $p \in \mathsf{Prop}$ y $i \in \mathsf{Agt}$. Las constantes y demás operadores booleanos se definen como $\top := (p \vee \neg p)$, $\bot := \neg\top$, $(\varphi \to \psi) := (\neg\varphi \vee \psi)$, $(\varphi \wedge \psi) := (\neg\varphi \vee \neg\psi)$. Las fórmulas de la forma $\mathsf{Kh}_i(\psi, \varphi)$ se leen como *"cuando ψ se cumple, el agente i sabe cómo hacer φ verdadera"*.

Definición 4.2 (LTS basado en Incertidumbre)**.** Un LTS *multiagente basado en incertidumbre* (LTS^U) para Prop, Act y Agt es una tupla $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \sim, \mathrm{V}\rangle$ donde $\langle \mathrm{S}, \mathrm{R}, \mathrm{V}\rangle$ es un LTS y $\sim$ asigna, a cada agente $i \in \mathsf{Agt}$, una relación de equivalencia de *indistinguibilidad*, denotada como $\sim_i$, sobre un conjunto no vacío de planes $\mathrm{P}_i \subseteq \mathsf{Act}^*$.

Intuitivamente, P_i es el conjunto de planes que el agente i tiene a su disposición. Es decir, los planes que percibe como posibles. Por otro lado, la representación gráfica de un LTS^U es similar a la de un LTS con el agregado de una relación de indistinguibilidad adyacente por cada agente considerado previamente.

Ejemplo 4.1. Dado el LTS $\mathfrak{L}$ del Ejemplo 3.4, el siguiente es un LTS^U $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \{\sim_i\}, \mathrm{V}\rangle$ con $\mathsf{Agt} = \{i\}$:



$$\sim_i = \left\{ \begin{array}{c} (a,a),\ (b,b),\ (ab,ab), \\ (ab,a),\ (a,ab) \end{array} \right\}$$

$$\mathrm{P}_i = \{a, b, ab\}$$

□

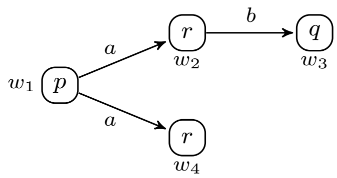
Análogamente a la lógica epistémica clásica [62], $\sim_i \subseteq \mathrm{P}_i \times \mathrm{P}_i$ describe la indistinguibilidad del agente i sobre los planes disponibles. Sin embargo, a lo largo de este trabajo se tomará una notación ligeramente diferente que hará más clara la comparación con la semántica basada en LTS, y simplificará algunas definiciones más adelante.

Observación 4.1. Sean $\langle \mathrm{S}, \mathrm{R}, \sim, \mathrm{V}\rangle$ un LTS^U, $i \in \mathsf{Agt}$ un agente y $\sigma \in \mathrm{P}_i$ un plan. Sean $[\sigma]_i$ la clase de equivalencia vía $\sim_i$ (es decir, $[\sigma]_i := \{\sigma' \in \mathrm{P}_i \mid \sigma \sim_i \sigma'\}$). Existe una correspondencia uno a uno entre cada $\sim_i$ y su conjunto de clases de equivalencias inducido, el cual se denota como una función $\mathrm{U} : \mathsf{Agt} \to 2^{2^{\mathsf{Act}^*}}$ tal que $\mathrm{U}(i) := \{[\sigma]_i \mid \sigma \in \mathrm{P}_i\}$. Por ende, a partir de ahora, un LTS^U será presentado como una tupla $\langle \mathrm{S}, \mathrm{R}, \{\mathrm{U}(i)\}_{i \in \mathsf{Agt}}, \mathrm{V}\rangle$, o en algunos casos como $\langle \mathrm{S}, \mathrm{R}, \mathrm{U}, \mathrm{V}\rangle$. Cada $\mathrm{U}(i)$ cumplirá con las siguientes propiedades:

- $\mathrm{U}(i) \neq \emptyset$ (dado que $\mathrm{P}_i \neq \emptyset$),
- si $\pi_1, \pi_2 \in \mathrm{U}(i)$ y $\pi_1 \neq \pi_2$, entonces $\pi_1 \cap \pi_2 = \emptyset$ (las clases de equivalencia son disjuntas de a pares),
- $\mathrm{P}_i = \bigcup_{\pi \in \mathrm{U}(i)} \pi$ (su unión es exactamente P_i), y
- $\emptyset \notin \mathrm{U}(i)$ (el conjunto vacío no es una clase de equivalencia).

Nuevamente se introduce la representación gráfica adaptada.

Ejemplo 4.2. Dado el LTS^U $\mathfrak{M}$ del Ejemplo 4.1, se tiene la siguiente la representación:



$$\mathrm{U}(i) = \left\{ \begin{array}{c} \{b\}, \\ \{ab, a\} \end{array} \right\}$$

□

Dada la incertidumbre del agente sobre Act^*, las habilidades de un agente i dependen no sólo de qué puede lograr un sólo plan, sino de qué puede lograr un conjunto de estos.

Definición 4.3. Sean $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \sim, \mathrm{V}\rangle$ un LTS^U, $\pi \subseteq \mathsf{Act}^*$, $u \in \mathrm{S}$ y $U \subseteq \mathrm{S}$, se define

$$\mathrm{R}_\pi := \bigcup_{\sigma \in \pi} \mathrm{R}_\sigma, \qquad \mathrm{R}_\pi(u) := \bigcup_{\sigma \in \pi} \mathrm{R}_\sigma(u), \qquad \mathrm{R}_\pi(U) := \bigcup_{u \in U} \mathrm{R}_\pi(u).$$

Ejemplo 4.3. Usando el LTS^U $\mathfrak{M}$ del Ejemplo 4.2, consideremos los siguientes conjuntos R_π para diferentes $\pi \subseteq \mathsf{Act}^*$ (notar que algunos de ellos no pertenecen necesariamente a P_i).

- $\mathrm{R}_{\{a,ab\}} = \mathrm{R}_a \cup \mathrm{R}_{ab} = \{(w_1, w_2), (w_1, w_4)\} \cup \{(w_1, w_3)\} = \{(w_1, w_2), (w_1, w_4), (w_1, w_3)\}$,
- $\mathrm{R}_{\{a,ab\}}(w_1) = \mathrm{R}_a(w_1) \cup \mathrm{R}_{ab}(w_1) = \{w_2, w_4\} \cup \{w_3\} = \{w_2, w_4, w_3\}$,
- $\mathrm{R}_{\{\epsilon,b\}} = \mathrm{R}_\epsilon \cup \mathrm{R}_b = \{(w_1, w_1), (w_2, w_2), (w_3, w_3), (w_4, w_4), (w_2, w_3)\}$, y
- $\mathrm{R}_{\{\epsilon,b\}}(w_2) = \mathrm{R}_\epsilon(w_2) \cup \mathrm{R}_b(w_2) = \{w_2\} \cup \{w_3\} = \{w_2, w_3\}$.

□

Con esto se puede generalizar la noción de ejecutabilidad fuerte para un conjunto de planes.

Definición 4.4 (Ejecutabilidad fuerte)**.** Sea $\{\mathrm{R}_a \subseteq \mathrm{S} \times \mathrm{S} \mid a \in A\}$ una colección de relaciones binarias. Un *conjunto de planes* $\pi \subseteq \mathsf{Act}^*$ es *fuertemente ejecutable* en $u \in \mathrm{S}$ si y sólo si *cada* plan $\sigma \in \pi$ es *fuertemente ejecutable* en u. Por lo tanto, $\mathrm{SE}(\pi) := \bigcap_{\sigma \in \pi} \mathrm{SE}(\sigma)$ es el conjunto de estados en S donde π es fuertemente ejecutable.

Ejemplo 4.4. Tomando el modelo del Ejemplo 4.2, por el Ejemplo 3.6 se sabe que $\mathrm{SE}(\epsilon) = \mathrm{S}$, $\mathrm{SE}(a) = \{w_1\}$, $\mathrm{SE}(b) = \{w_2\}$ y $\mathrm{SE}(\sigma) = \emptyset$ para todo $\sigma \notin \{\epsilon, a, b\}$. Luego $\mathrm{SE}(\{a, ab\}) = \mathrm{SE}(a) \cap \mathrm{SE}(ab) = \{w_1\} \cap \emptyset = \emptyset$ y $\mathrm{SE}(\{\epsilon, b\}) = \mathrm{SE}(\epsilon) \cap \mathrm{SE}(b) = \mathrm{S} \cap \{w_2\} = \{w_2\}$. □

Una vez introducidos los modelos y definiciones asociadas, se procede a presentar la semántica de la lógica $\mathsf{L}_{\mathsf{Kh}_i}$.

Definición 4.5 ($\mathsf{L}_{\mathsf{Kh}_i}$ en LTS^{U}s)**.** Sean $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \{\mathrm{U}(i)\}_{i\in\mathsf{Agt}}, \mathrm{V}\rangle$ un LTS^U sobre Act, Prop y Agt, y $w \in \mathrm{S}$. La relación de satisfacibilidad $\models$ entre $\mathfrak{M}$, w y las fórmulas de $\mathsf{L}_{\mathsf{Kh}_i}$ se define inductivamente de la siguiente manera:

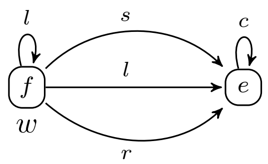
$$
\begin{array}{lll}
\mathfrak{M}, w \models p & \text{sii} & p \in \mathrm{V}(w), \\
\mathfrak{M}, w \models \neg\varphi & \text{sii} & \mathfrak{M}, w \not\models \varphi, \\
\mathfrak{M}, w \models \varphi \vee \psi & \text{sii} & \mathfrak{M}, w \models \varphi \text{ o } \mathfrak{M}, w \models \psi, \\
\mathfrak{M}, w \models \mathsf{Kh}_i(\psi, \varphi) & \text{sii} & \text{existe } \pi \in \mathrm{U}(i) \text{ tal que} \\
 & & \textbf{(Kh-1) } [\![\psi]\!]^{\mathfrak{M}} \subseteq \mathrm{SE}(\pi) \text{ y } \textbf{(Kh-2) } \mathrm{R}_\pi([\![\psi]\!]^{\mathfrak{M}}) \subseteq [\![\varphi]\!]^{\mathfrak{M}},
\end{array}
$$

con $[\![\chi]\!]^{\mathfrak{M}} := \{w \in \mathrm{S} \mid \mathfrak{M}, w \models \chi\}$. Una fórmula φ es satisfacible sii $\mathfrak{M}, w \models \varphi$ para algún $\mathfrak{M}$ y $w \in \mathrm{D}_{\mathfrak{M}}$, y es válida (denotado como $\models \varphi$) sii $\mathfrak{M}, w \models \varphi$ para todo $\mathfrak{M}$ y $w \in \mathrm{D}_{\mathfrak{M}}$. Una fórmula φ se cumple (globalmente) en un LTS^U $\mathfrak{M}$ ($\mathfrak{M} \models \varphi$) sii $\mathfrak{M}, w \models \varphi$ para todo $w \in \mathrm{D}_{\mathfrak{M}}$. Para $\Psi \subseteq \mathsf{L}_{\mathsf{Kh}}$, $\mathfrak{M}, w \models \Psi$ sii $\mathfrak{M}, w \models \psi$ para todo $\psi \in \Psi$; y $\Psi \models \varphi$ sii $\mathfrak{M}, w \models \Psi$ implica $\mathfrak{M}, w \models \varphi$ para todo $\mathfrak{M}$ y $w \in \mathrm{D}_{\mathfrak{M}}$. El conjunto de planes π que hace verdadero $\mathsf{Kh}_i(\psi, \varphi)$ en $\mathfrak{M}$ es llamado *testigo* para $\mathsf{Kh}_i(\psi, \varphi)$ en $\mathfrak{M}$.

Ejemplo 4.5. Dado el modelo del Ejemplo 4.2 con $\mathrm{U}(i) = \{\{b\}, \{ab, a\}\}$. Como sucedía con el Ejemplo 3.7, $\mathfrak{M}, w_1 \models p$. Como $\mathrm{SE}(\{b\}) = \{w_2\}$ y $\mathrm{SE}(\{ab, a\}) = \emptyset$, entonces no existe $\pi \in \mathrm{U}(i)$ tal que $[\![p]\!]^{\mathfrak{M}} = \{w_1\} \subseteq \mathrm{SE}(\pi)$. Por lo tanto, $\mathfrak{M}, w_1 \not\models \mathsf{Kh}(p, q)$ y más aún $\mathfrak{M}, w_1 \not\models \mathsf{Kh}(p, r)$. □

Comparando la Definición 3.7 con la Definición 4.5, $\mathsf{Kh}_i(\psi, \varphi)$ sigue actuando *globalmente*. La diferencia es que ahora se necesita que el *agente* i tenga un *conjunto de planes* que satisfagan ejecutabilidad fuerte en cada estado ψ (condición **(Kh-1)**). Y que además dicho conjunto debe funcionar como pasaba con el plan testigo anteriormente: al ser ejecutado en estados ψ, debe terminar, sin fallas, en estados satisfaciendo φ (condición **(Kh-2)**). El siguiente ejemplo ilustra de una manera más integral los conceptos explicados anteriormente.

Ejemplo 4.6. Consideremos un protocolo de evacuación de un edificio. Naturalmente, existen ciertos cursos de acción que, en caso de emergencia, llevan a un punto de reunión seguro y otros que no. Por ejemplo, si un incendio (f) ocurre, hay tres posibles rutas de escape: usar las escaleras (acción s), usar la rampa (acción r), o usar el ascensor (acción l). Usar las escaleras o la rampa garantiza que los agentes llegarán al punto de reunión (e). Esto no sucede cuando se usa el ascensor ya que durante un incendio puede cortarse la electricidad y el agente puede quedar encerrado. El protocolo de evacuación indica que, en caso de incendio, los agentes deben tomar las escaleras o la rampa y llamar al 100 (acción c). Por lo tanto, los planes a usar en caso de incendio son sc y rc. Mientras tanto, tomar el ascensor está permitido. Esta situación puede ilustrarse con el siguiente LTS^U $\mathfrak{M}$:



$\mathrm{U}(i) = \{\{sc, rc\}, \{lc\}\}$
$\mathrm{U}(j) = \{\{lc, sc, rc\}\}$

A la derecha del LTS, los conjuntos $\mathrm{U}(i)$ y $\mathrm{U}(j)$ muestran la incertidumbre de los agentes i y j. Ambos agentes son conscientes de la existencia de tres cursos de acción posibles, es decir, $\{sc, rc, lc\} \subseteq \mathrm{P}_i \cap \mathrm{P}_j$. El agente i, que ha tomado un curso de emergencias, considera que sc y rc son estrategias de evacuación igualmente buenas en caso de incendio, mientras que lc no. De hecho, siguiendo la Definición 4.5, se tiene que $\mathfrak{M}, w \models \mathsf{Kh}_i(f, e)$. Es decir, el agente i sabe cómo conseguir e dado f, ya que el conjunto $\{sc, rc\}$ es un testigo de la formula. Por otro lado, el agente j tiene una sola clase de equivalencia, y uno de sus planes (lc) no es fuertemente ejecutable en el estado donde vale f. Por lo tanto, $\mathfrak{M}, w \not\models \mathsf{Kh}_j(f, e)$. □

En la semántica no hay restricciones en cuanto a la incertidumbre de los $\mathrm{U}(i)$. Se pueden considerar algunas posibilidades naturales como por ejemplo, que $a \sim_i b$ implique que $ab \sim_i ba$. Es decir, que la indistinguibilidad atómica implique indistinguibilidad al nivel de planes más complejos. Sin embargo, consideremos la siguiente situación en la que esta propiedad no se mantiene. Supongamos que un agente i usualmente tiene hambre a la tarde y suele comer una colación. A veces es algo dulce, otras veces es salado, pero no tiene ninguna preferencia entre las dos (es decir, $a \sim_i b$). Cuando está más hambriento, usualmente come los dos (algo salado y algo dulce). Pero en esos casos prefiere comer primero lo salado y luego lo dulce (como postre). Jamás lo haría al revés (es decir, no sucede que $ab \sim_i ba$). Esta lógica pretende dar una base para modelar el saber cómo de los agentes y, por lo tanto, es más conveniente sólo imponer condiciones mínimas para la relación de indistinguibilidad.

Volviendo a las propiedades semánticas de la lógica, así como sucedía en la Proposición 3.1, la modalidad universal A también es definible en $\mathsf{L}_{\mathsf{Kh}_i}$ sobre LTS^U. Esto es posible ya que $\mathrm{U}(i) \neq \emptyset$ y $\emptyset \notin \mathrm{U}(i)$, mencionado en la Observación 4.1.

Proposición 4.1. *Sean $\mathfrak{M}$ un LTS^U y $w \in \mathrm{D}_{\mathfrak{M}}$. Entonces,*

$$\textit{existe } i \in \mathsf{Agt} \textit{ tal que } \mathfrak{M}, w \models \mathsf{Kh}_i(\neg\varphi, \bot) \qquad \textit{sii} \qquad [\![\varphi]\!]^{\mathfrak{M}} = \mathrm{D}_{\mathfrak{M}}.$$

Demostración. ($\Rightarrow$) Sea $i \in \mathsf{Agt}$ un agente tal que $\mathfrak{M}, w \models \mathsf{Kh}_i(\neg\varphi, \bot)$. Luego existe $\pi \in \mathrm{U}(i)$ tal que **(Kh-1)** $[\![\neg\varphi]\!]^{\mathfrak{M}} \subseteq \mathrm{SE}(\pi)$ y **(Kh-2)** $\mathrm{R}_\pi([\![\neg\varphi]\!]^{\mathfrak{M}}) \subseteq [\![\bot]\!]^{\mathfrak{M}}$. Si $[\![\varphi]\!]^{\mathfrak{M}} \neq \mathrm{D}_{\mathfrak{M}}$, entonces existe $u \in [\![\neg\varphi]\!]^{\mathfrak{M}}$. Por lo tanto, **(Kh-1)** implica $u \in \mathrm{SE}(\pi) = \bigcap_{\sigma\in\pi} \mathrm{SE}(\sigma)$. Como $\pi \in \mathrm{U}(i)$, $\pi \neq \emptyset$ y existe $\sigma \in \pi$ con $u \in \mathrm{SE}(\sigma)$. Por ende, $\mathrm{R}_\sigma(u) \neq \emptyset$, $\mathrm{R}_\pi(u) \neq \emptyset$ y con esto $\mathrm{R}_\pi([\![\neg\varphi]\!]^{\mathfrak{M}}) \neq \emptyset$, que sería, $\emptyset \subsetneq \mathrm{R}_\pi([\![\neg\varphi]\!]^{\mathfrak{M}})$. Pero luego, de **(Kh-2)**, $\emptyset \subsetneq \mathrm{R}_\pi([\![\neg\varphi]\!]^{\mathfrak{M}}) \subseteq [\![\bot]\!]^{\mathfrak{M}}$, es decir, $\emptyset \subsetneq [\![\bot]\!]^{\mathfrak{M}}$, una contradicción. Por ende, $[\![\varphi]\!]^{\mathfrak{M}} = \mathrm{D}_{\mathfrak{M}}$.

($\Leftarrow$) Si $[\![\varphi]\!]^{\mathfrak{M}} = \mathrm{D}_{\mathfrak{M}}$, entonces $[\![\neg\varphi]\!]^{\mathfrak{M}} = \emptyset$ y **(Kh-1)** en la definición semántica $\mathsf{Kh}_i(\neg\varphi, \bot)$ se cumple para todo $\pi \in 2^{\mathsf{Act}^*}$. Más aún, $\mathrm{R}_\pi([\![\neg\varphi]\!]^{\mathfrak{M}}) = \bigcup_{u\in[\![\neg\varphi]\!]^{\mathfrak{M}}} \mathrm{R}_\pi(u) = \bigcup_{u\in\emptyset} \mathrm{R}_\pi(u) = \emptyset$, así que **(Kh-2)** también se cumple para cualquier π. Finalmente, $\mathrm{U}(i) \neq \emptyset$ (así que existe $\pi \in \mathrm{U}(i)$) y $\mathsf{Agt} \neq \emptyset$ (existe un agente $i \in \mathsf{Agt}$) Por ende, para un agente $i \in \mathsf{Agt}$ se tiene que $\mathfrak{M}, w \models \mathsf{Kh}_i(\neg\varphi, \bot)$. $\square$

La elección de $\pi \in \mathrm{U}(i)$ en la demostración es independiente del agente i considerado, ya que sólo se necesita que $\mathrm{U}(i) \neq \emptyset$ y $\emptyset \notin \mathrm{U}(i)$, una propiedad que se cumple para todos los agentes involucrados. Con esto, se puede deducir el mismo resultado de la Proposición 4.1 cambiando el "existe" por un "para todo".

Proposición 4.2. *Sean $\mathfrak{M}$ un LTS^U y $w \in \mathrm{D}_{\mathfrak{M}}$.*

$$\textit{para todo } i \in \mathsf{Agt} \textit{ tal que } \mathfrak{M}, w \models \mathsf{Kh}_i(\neg\varphi, \bot) \qquad \textit{sii} \qquad [\![\varphi]\!]^{\mathfrak{M}} = \mathrm{D}_{\mathfrak{M}}.$$

Más aún, con un razonamiento análogo, se puede conseguir la misma propiedad fijando para un solo agente.

Proposición 4.3. *Sean $\mathfrak{M}$ un LTS^U y $w \in \mathrm{D}_{\mathfrak{M}}$. Entonces, para cualquier $j \in \mathsf{Agt}$*

$$\mathfrak{M}, w \models \mathsf{Kh}_j(\neg\varphi, \bot) \qquad \textit{sii} \qquad [\![\varphi]\!]^{\mathfrak{M}} = \mathrm{D}_{\mathfrak{M}}.$$

Por lo tanto, la modalidad universal se puede definir de la forma $\mathsf{A}\varphi := \bigvee_{i\in\mathsf{Agt}} \mathsf{Kh}_i(\neg\varphi, \bot)$ ó $\mathsf{A}\varphi := \bigwedge_{i\in\mathsf{Agt}} \mathsf{Kh}_i(\neg\varphi, \bot)$ dado que Agt es un conjunto finito no vacío. O definirlo para un agente fijo j como $\mathsf{A}\varphi := \mathsf{Kh}_j(\neg\varphi, \bot)$. De cualquier manera, las tres definiciones de A son equivalentes y son tres formas de generalizar la modalidad A de L_{Kh}. Para este trabajo se tomará la alternativa con $\bigvee$. Luego se define $\mathsf{E}\varphi := \neg\mathsf{A}\neg\varphi$.

Ejemplo 4.7. Teniendo en cuenta el Ejemplo 4.2 y el mismo razonamiento en el Ejemplo 3.8, se puede concluir que $\mathfrak{M}, w_1 \models \mathsf{A}(p \vee q \vee r)$ tomando cualquier $\pi \in \mathrm{U}(i)$ (dado que $\mathrm{U}(i) \neq \emptyset$ y $\emptyset \notin \mathrm{U}(i)$). $\square$

Ahora bien, claramente diferentes agentes tienen diferente consciencia sobre sus propias habilidades. Al mismo tiempo, debido a la naturaleza global del operador de saber cómo, se cumple que

$$\mathfrak{M}, w \models \mathsf{Kh}_i(\psi, \varphi) \text{ si y sólo si } \mathfrak{M}, w \models \mathsf{AKh}_i(\psi, \varphi),$$

o equivalentemente,

$$\mathfrak{M}, w \models \mathsf{Kh}_j(\neg\mathsf{Kh}_i(\psi, \varphi), \bot), \text{ para algún agente } j.$$

Pero esto no implica que los agentes *saben que* "el agente i sabe cómo cumplir φ dado ψ". Lo que sucede en esta situación es que $\mathsf{Kh}_i(\psi, \varphi)$ se vuelve objetivamente verdadera, y que asumir su negación conlleva a una contradicción. No hay una noción de incertidumbre sobre estados en estos modelos que pudieran conllevar a una noción de "saber que".

Por último, se puede argumentar que dada la noción de indistinguibilidad epistémica entre planes, un agente debe saber que un determinado plan es distinguible o no de otro, o que un agente es consciente de la disponibilidad de un curso de acción en particular. Sin embargo, las modalidades de saber cómo no pueden hablar de la relación en sí misma, sino de la existencia de un conjunto de planes indistinguibles entre sí y de las consecuencias de ejecutar estos.

Capítulo 5

Axiomatización y completitud

> A proof tells us where to concentrate our doubts.
>
> Morris Kline

En este capítulo nos dedicaremos a estudiar un aspecto fundamental a la hora de comprender y caracterizar una lógica. Introduciremos un sistema axiomático a la Hilbert para la lógica $\mathsf{L}_{\mathsf{Kh}_i}$, trabajada en [8, 10]. Como es usual, dicho sistema consiste de un conjunto de axiomas y reglas de inferencia que nos permiten demostrar teoremas o, de manera más general, deducir fórmulas a partir de otras mediante una noción de consecuencia sintáctiva. Dicha noción de consecuencia sintáctica (denotada como $\vdash$) resulta adecuada como caracterización de la lógica si coincide exactamente con la noción de consecuencia semántica (denotada como $\models$), es decir que captura las deducciones que pueden realizarse a nivel de modelos. Cuando esto ocurre, diremos que el sistema es correcto y fuertemente completo para $\mathsf{L}_{\mathsf{Kh}_i}$.

Vale la pena recordar que en el Capítulo 3, se pusieron bajo discusión dos axiomas de la Tabla 3.1: EMP y COMPKh. El argumento de dicha discusión es que consideramos que ambos implican, a nivel conceptual, cierto nivel de omnisciencia [34, 54, 84, 102, 112]. En nuestro sistema, dichos axiomas no aparecen. Más aún, ambas fórmulas no son válidas (y por lo tanto no deducibles) en nuestra lógica. De esta manera, demostramos que conceptualmente la lógica se comporta de la manera buscada.

5.1. El sistema axiomático $\mathcal{L}_{\mathsf{Kh}_i}^{\mathrm{LTS}^U}$

En esta sección introduciremos los axiomas y las reglas del sistema que llamaremos $\mathcal{L}_{\mathsf{Kh}_i}^{\mathrm{LTS}^U}$. Luego demostraremos que los axiomas EMP y COMPKh de [116–118] no son admisibles en la semántica basada en LTS^U. Finalmente, introduciremos otras definiciones necesarias para el resto de la tesis, como los conceptos de derivabilidad, consistencia, maximalidad, modelo canónico, entre otros.

La siguiente tabla (Tabla 5.1) es el sistema axiomático $\mathcal{L}_{\mathsf{Kh}_i}^{\mathrm{LTS}^U}$ de $\mathsf{L}_{\mathsf{Kh}_i}$ sobre LTS^U. El primer bloque, $\mathcal{L}$, y las reglas son las mismas que se encuentran en la Tabla 3.1 reemplazando Kh por Kh_i. Esto se debe a que, como la modalidad universal A es definible también con Kh_i, las formulas y reglas en $\mathcal{L}$ siguen cumpliéndose en la semántica basada en LTS^U. El axioma del segundo bloque $\mathcal{L}_{\mathrm{LTS}^U}$, KhA, es más débil comparado con el segundo bloque de la Tabla 3.1 y es derivable de $\mathcal{L}_{\mathsf{Kh}}^{\mathrm{LTS}}$ (ver la Proposición 6.2). Sin embargo, el mismo puede ser discutible con los argumentos usados para EMP y COMPKh sobre la omnisciencia del agente. KhA describe que agente todavía puede transformar información factual, provista por las implicaciones globales, y transformarlas en conocimiento nuevo, aunque en un nivel menos idealizado. Se puede verificar, además, que los axiomas EMP y COMPKh de $\mathcal{L}_{\mathsf{Kh}}^{\mathrm{LTS}}$ no son válidos en los modelos LTS^U.

Ejemplo 5.1. Consideremos un LTS^U $\mathfrak{M}$ con un agente i tal que $\mathrm{U}(i) = \{\{a\}, \{b\}\}$:

$$\mathfrak{M}: \quad \underset{w_1}{\boxed{p}} \xrightarrow{a} \underset{w_2}{\boxed{q}} \xrightarrow{b} \underset{w_3}{\boxed{r}}$$

Bloque $\mathcal{L}$:	TAUT	$\vdash \varphi$ para φ una tautología proposicional
	DISTA	$\vdash \mathsf{A}(\varphi \to \psi) \to (\mathsf{A}\varphi \to \mathsf{A}\psi)$
	TA	$\vdash \mathsf{A}\varphi \to \varphi$
	$\mathsf{4KhA}$	$\vdash \mathsf{Kh}_i(\psi, \varphi) \to \mathsf{AKh}_i(\psi, \varphi)$
	$\mathsf{5KhA}$	$\vdash \neg\mathsf{Kh}_i(\psi, \varphi) \to \mathsf{A}\neg\mathsf{Kh}_i(\psi, \varphi)$
Bloque $\mathcal{L}_{\mathrm{LTS}^U}$:	KhA	$\vdash (\mathsf{A}(\chi \to \psi) \land \mathsf{Kh}_i(\psi, \varphi) \land \mathsf{A}(\varphi \to \theta)) \to \mathsf{Kh}_i(\chi, \theta)$
Reglas:	$\dfrac{\varphi \quad (\varphi \to \psi)}{\psi}$ MP	$\dfrac{\varphi}{\mathsf{A}\varphi}$ NECA

Tabla 5.1: Axiomatización $\mathcal{L}_{\mathsf{Kh}_i}^{\mathrm{LTS}^U}$ de $\mathsf{L}_{\mathsf{Kh}_i}$ con respecto a los LTS^Us.

Si bien $\mathsf{A}(p \to p)$ se cumple, no es el caso para $\mathsf{Kh}_i(p, p)$ dado que el agente no tiene a su disposición un conjunto de planes testigo que lo lleve al mismo estado w_1 (por ejemplo, ϵ). Lo mismo sucede para $\mathsf{Kh}_i(q, q)$ y $\mathsf{Kh}_i(r, r)$. Por otro lado, $\mathsf{Kh}_i(p, q)$ y $\mathsf{Kh}_i(q, r)$ sí se cumplen, pero $\mathsf{Kh}_i(p, r)$ no ya que no hay un conjunto de planes disponible para el agente que "represente" una composición de los conjuntos $\{a\}$ y $\{b\}$. □

Para demostrar que el sistema axiomático es correcto (todo teorema es una tautología) y fuertemente completo (toda tautología es un teorema), es necesario dar antes una definición formal de cuándo una fórmula es "deducible" de un conjunto de supuestos en este sistema axiomático y cuándo un conjunto es "consistente" con respecto a este sistema.

Definición 5.1. Una fórmula φ de $\mathsf{L}_{\mathsf{Kh}_i}$ es un teorema en $\mathcal{L}_{\mathsf{Kh}_i}^{\mathrm{LTS}^U}$, $\vdash \varphi$, si es una instancia de un axioma o si es consecuencia de aplicar las reglas MP ó NECA. Sea Γ un conjunto de fórmulas de $\mathsf{L}_{\mathsf{Kh}_i}$, φ se deduce de Γ en $\mathcal{L}_{\mathsf{Kh}_i}^{\mathrm{LTS}^U}$, $\Gamma \vdash \varphi$, si y sólo si existen fórmulas $\varphi_1, \ldots, \varphi_n \in \Gamma$ tales que $\vdash (\varphi_1 \land \cdots \land \varphi_n) \to \varphi$. Caso contrario, se escribe $\Gamma \nvdash \varphi$. Γ es consistente con $\mathcal{L}_{\mathsf{Kh}_i}^{\mathrm{LTS}^U}$ si $\Gamma \nvdash \bot$. Si $\Gamma \vdash \bot$, Γ es inconsistente con $\mathcal{L}_{\mathsf{Kh}_i}^{\mathrm{LTS}^U}$.

Ejemplo 5.2. La instancia de $\mathsf{4KhA}$, con $\psi = p$ y $\varphi = q$, es un teorema. Por lo tanto, $\vdash \mathsf{Kh}_i(p, q) \to \mathsf{AKh}_i(p, q)$. Más aún, $\{\mathsf{Kh}_i(p, q)\} \vdash \mathsf{AKh}_i(p, q)$. Como $(r \land \neg r) = \bot$, usando una instancia de TAUT, $\varphi = (\psi \to \psi)$ con $\psi = (r \land \neg r)$, $\vdash (r \land \neg r) \to \bot$. Con esto, $\{r, \neg r\} \vdash \bot$ y $\{r, \neg r\}$ es inconsistente. □

Teniendo en cuenta estas definiciones, del sistema $\mathcal{L}_{\mathsf{Kh}_i}^{\mathrm{LTS}^U}$ se pueden deducir los siguientes teoremas, basados en [116, 118]. Tener en cuenta que estos serán importantes en este capítulo.

Proposición 5.1. *Las fórmulas* $(\mathsf{E}\psi \land \mathsf{Kh}_i(\psi, \varphi)) \to \mathsf{E}\varphi$ *(llamada KhE),* $\mathsf{A}\neg\psi \to \mathsf{Kh}_i(\psi, \varphi)$ *(llamada SCOND) y* $\mathsf{Kh}_i(\bot, \varphi)$ *(llamada COND) son derivables de* $\mathcal{L}_{\mathsf{Kh}_i}^{\mathrm{LTS}^U}$.
Es decir, ***(1)*** $\vdash (\mathsf{E}\psi \land \mathsf{Kh}_i(\psi, \varphi)) \to \mathsf{E}\varphi$ ***(2)*** $\vdash \mathsf{A}\neg\psi \to \mathsf{Kh}_i(\psi, \varphi)$ *y* ***(3)*** $\vdash \mathsf{Kh}_i(\bot, \varphi)$.

Demostración.

(KhE) Por TAUT, esto es equivalente a demostrar que

$$\vdash (\mathsf{E}\psi \land \mathsf{Kh}_i(\psi, \varphi) \land \mathsf{A}\neg\varphi) \to \bot. \tag{5.1}$$

Para ello, estableceremos una serie de implicaciones y, utilizando la transitividad de la implicación, se tendrá la propiedad de arriba. Aplicando TAUT, se tiene que

$$\vdash (\mathsf{E}\psi \land \mathsf{Kh}_i(\psi, \varphi) \land \mathsf{A}\neg\varphi) \to (\mathsf{E}\psi \land \mathsf{Kh}_i(\psi, \varphi) \land \mathsf{A}\neg\varphi).$$

Tomando TAUT ($\vdash \neg\varphi \leftrightarrow (\varphi \to \bot)$ y $\vdash \psi \to \psi$) y NECA, en el lado derecho se sustituye $\mathsf{A}\neg\varphi$ por $\mathsf{A}(\varphi \to \bot)$ y se añade $\mathsf{A}(\psi \to \psi)$, que es una tautología. Por lo tanto,

$$\vdash (\mathsf{E}\psi \land \mathsf{Kh}_i(\psi, \varphi) \land \mathsf{A}\neg\varphi) \to (\mathsf{E}\psi \land \mathsf{A}(\psi \to \psi) \land \mathsf{Kh}_i(\psi, \varphi) \land \mathsf{A}(\varphi \to \bot)). \tag{5.2}$$

Por otro lado, usando una instancia de KhA, se tiene que

$$\vdash (\mathsf{A}(\psi \to \psi) \wedge \mathsf{Kh}_i(\psi, \varphi) \wedge \mathsf{A}(\varphi \to \bot)) \to (\mathsf{Kh}_i(\psi, \bot)).$$

Agregando $\mathsf{E}\psi$ a ambos lados (usando una instancia de TAUT),

$$\vdash (\mathsf{E}\psi \wedge \mathsf{A}(\psi \to \psi) \wedge \mathsf{Kh}_i(\psi, \varphi) \wedge \mathsf{A}(\varphi \to \bot)) \to (\mathsf{E}\psi \wedge \mathsf{Kh}_i(\psi, \bot)).$$

Por definición de E y A,

$$\vdash (\mathsf{E}\psi \wedge \mathsf{A}(\psi \to \psi) \wedge \mathsf{Kh}_i(\psi, \varphi) \wedge \mathsf{A}(\varphi \to \bot)) \to (\neg\mathsf{A}\neg\psi \wedge \mathsf{A}\neg\psi).$$

Y por definición de $\bot$,

$$\vdash (\mathsf{E}\psi \wedge \mathsf{A}(\psi \to \psi) \wedge \mathsf{Kh}_i(\psi, \varphi) \wedge \mathsf{A}(\varphi \to \bot)) \to \bot. \tag{5.3}$$

Dadas las implicaciones 5.2 y 5.3, por transitividad se obtiene 5.1.

(SCOND) Aplicando una instancia de KhA,

$$(\mathsf{A}(\psi \to \psi) \wedge \mathsf{Kh}_i(\psi, \bot) \wedge \mathsf{A}(\bot \to \varphi)) \to \mathsf{Kh}_i(\psi, \varphi).$$

Por TAUT, esto equivale a

$$\vdash \mathsf{A}(\psi \to \psi) \to (\mathsf{A}(\bot \to \varphi) \to (\mathsf{Kh}_i(\psi, \bot) \to \mathsf{Kh}_i(\psi, \varphi))).$$

Usando TAUT y NECA se tiene que $\vdash \mathsf{A}(\psi \to \psi)$ y $\vdash \mathsf{A}(\bot \to \varphi)$. Luego, usando MP dos veces $\vdash \mathsf{Kh}_i(\psi, \bot) \to \mathsf{Kh}_i(\psi, \varphi)$. Por definición de A (teniendo en cuenta la Proposición 4.1), $\vdash \mathsf{A}\neg\psi \to \mathsf{Kh}_i(\psi, \varphi)$.

(COND) Sea $\vdash \mathsf{A}\neg\bot \to \mathsf{Kh}_i(\bot, \varphi)$ una instancia del item anterior. Usando TAUT y NECA se tiene que $\vdash \mathsf{A}\neg\bot$. Por MP, $\vdash \mathsf{Kh}_i(\bot, \varphi)$.

□

Una diferencia importante del sistema de la Tabla 5.1 del visto en [8, 10] es que KhE originalmente era un axioma. Sin embargo, durante la escritura de este trabajo, se dio con una demostración de que era deducible de KhA. Esto beneficia el sistema actual en su simplicidad, pues es el resultado de reemplazar dos axiomas (EMP y COMPKh) de la propuesta original, por uno solo (KhA).

Con esto, estamos en posición de dar una formalización de correctitud y completitud fuerte para esta lógica. El sistema $\mathcal{L}_{\mathsf{Kh}_i}^{\mathrm{LTS}^U}$ es correcto si para todo conjunto de fórmulas $\Gamma \cup \{\varphi\}$ de $\mathsf{L}_{\mathsf{Kh}_i}$, $\Gamma \vdash \varphi$ implica $\Gamma \models \varphi$; y es fuertemente completo si para todo conjunto de fórmulas $\Gamma \cup \{\varphi\}$ de $\mathsf{L}_{\mathsf{Kh}_i}$, $\Gamma \models \varphi$ implica $\Gamma \vdash \varphi$. Por un lado, demostrar correctitud se reduce a demostrar que los axiomas y reglas son válidos, un resultado que termina siendo directo. Por el otro, completitud fuerte requiere un poco más de trabajo. En especial, la implicación anterior se puede reformular en términos de consistencia. "$\Gamma \models \varphi$ implica $\Gamma \vdash \varphi$" es equivalente a que "si $\Gamma \cup \{\neg\varphi\}$ es consistente con $\mathcal{L}_{\mathsf{Kh}_i}^{\mathrm{LTS}^U}$, entonces $\Gamma \cup \{\neg\varphi\}$ es satisfacible". Esto se puede simplificar a que $\mathcal{L}_{\mathsf{Kh}_i}^{\mathrm{LTS}^U}$ es fuertemente completo si y sólo si para todo conjunto de fórmulas Γ de $\mathsf{L}_{\mathsf{Kh}_i}$, si Γ es consistente, entonces Γ es satisfacible. Por estos motivos, en la siguiente sección, y por consiguiente en el resto del capítulo, nos enfocaremos en la prueba de esta propiedad.

5.2. Completitud

La estrategia para demostrar completitud fuerte para $\mathcal{L}_{\mathsf{Kh}_i}^{\mathrm{LTS}^U}$ se basa en [24, Proposición 4.12] y es la de, dado un Γ consistente, construir con el sistema $\mathcal{L}_{\mathsf{Kh}_i}^{\mathrm{LTS}^U}$ un LTS^U que lo satisfaga. Para esto será importante definir el concepto de conjuntos maximales consistentes, que son conjuntos consistentes tales que al agregarle una fórmula nueva, terminan siendo inconsistentes. La definición siguiente se basa en [24].

Definición 5.2. Sea Γ un conjunto de fórmulas de $\mathsf{L}_{\mathsf{Kh}_i}$, Γ es un conjunto maximal consistente con $\mathcal{L}_{\mathsf{Kh}_i}^{\mathrm{LTS}^U}$ si Γ es consistente con $\mathcal{L}_{\mathsf{Kh}_i}^{\mathrm{LTS}^U}$ y cualquier conjunto Δ que contenga propiamente a Γ ($\Gamma \subsetneq \Delta$) es inconsistente con $\mathcal{L}_{\mathsf{Kh}_i}^{\mathrm{LTS}^U}$.

Ejemplo 5.3. El conjunto $\Gamma = \mathsf{Prop}$ es consistente dado que no existen $p_1, \ldots, p_n \in \Gamma$ tales que $\vdash (p_1 \wedge \cdots \wedge p_n) \rightarrow \bot$. En cambio, el conjunto $\Gamma' = \mathsf{Prop} \cup \{p \rightarrow \neg q\}$ (con $p, q \in \mathsf{Prop}$) es inconsistente ya que $\vdash ((p \rightarrow \neg q) \wedge p \wedge q) \rightarrow \bot$. Γ es consistente pero no maximal consistente ya que $\Gamma \cup \{p \rightarrow p\}$ es consistente y $\Gamma \subsetneq \Gamma \cup \{p \rightarrow p\}$. □

Las siguientes propiedades para los conjuntos maximales consistentes, basadas en [24, Proposición 4.16], también nos serán útiles.

Observación 5.1. Dado Γ un conjunto maximal consistente con $\mathcal{L}_{\mathsf{Kh}_i}^{\mathrm{LTS}^U}$:

- para toda fórmula φ en $\mathsf{L}_{\mathsf{Kh}_i}$, $\varphi \in \Gamma$ si y sólo si $\Gamma \vdash \varphi$;
- para todo par de fórmulas φ, ψ en $\mathsf{L}_{\mathsf{Kh}_i}$, si $\varphi, \varphi \rightarrow \psi \in \Gamma$ entonces $\psi \in \Gamma$; y
- para toda fórmula φ en $\mathsf{L}_{\mathsf{Kh}_i}$, $\varphi \in \Gamma$ ó $\neg\varphi \in \Gamma$.

Más aún, dado que el conjunto de fórmulas de L_{Kh} es numerable, todo conjunto consistente con $\mathcal{L}_{\mathsf{Kh}_i}^{\mathrm{LTS}^U}$ puede extenderse a un conjunto maximal consistente, mediante una construcción de Lindenbaum estándar como la vista en [24, Lema 4.17].

Observación 5.2. Todo conjunto Γ consistente con $\mathcal{L}_{\mathsf{Kh}_i}^{\mathrm{LTS}^U}$ puede extenderse a un conjunto Γ' maximal consistente con $\mathcal{L}_{\mathsf{Kh}_i}^{\mathrm{LTS}^U}$.

Con esta última observación, la demostración de completitud fuerte se reduce a, dado un Γ consistente, extenderlo a un conjunto maximal consistente Γ' y construir a partir del mismo y con $\mathcal{L}_{\mathsf{Kh}_i}^{\mathrm{LTS}^U}$, un modelo LTS^U (llamado canónico) que satisfaga Γ.

Definición 5.3. Sea $\mathbf{\Phi}$ el conjunto de todos los conjuntos de fórmulas en $\mathsf{L}_{\mathsf{Kh}_i}$ maximales consistentes con $\mathcal{L}_{\mathsf{Kh}_i}^{\mathrm{LTS}^U}$. Para cualquier $\Delta \in \mathbf{\Phi}$, se definen

$$\begin{array}{llll} \Delta|_{\mathsf{Kh}_i} := \{\mathsf{Kh}_i(\psi,\varphi) \mid \mathsf{Kh}_i(\psi,\varphi) \in \Delta\}, & & \Delta|_{\mathsf{Kh}} := \bigcup_{i \in \mathsf{Agt}} \Delta|_{\mathsf{Kh}_i}. \\ \Delta|_{\neg\mathsf{Kh}_i} := \{\neg\mathsf{Kh}_i(\psi,\varphi) \mid \neg\mathsf{Kh}_i(\psi,\varphi) \in \Delta\}, & & \Delta|_{\neg\mathsf{Kh}} := \bigcup_{i \in \mathsf{Agt}} \Delta|_{\neg\mathsf{Kh}_i}. \end{array}$$

Sea Γ un conjunto en $\mathbf{\Phi}$; se define una estructura que satisface sus fórmulas. Sea el conjunto de acciones básicas $\mathsf{Act}_i^\Gamma := \{\langle\psi,\varphi\rangle \mid \mathsf{Kh}_i(\psi,\varphi) \in \Gamma\}$ asociado a cada agente $i \in \mathsf{Agt}$, y su unión $\mathsf{Act}^\Gamma := \bigcup_{i \in \mathsf{Agt}} \mathsf{Act}_i^\Gamma$. Notar que $\mathsf{Kh}_i(\bot,\varphi) \in \Gamma$ para todo $i \in \mathsf{Agt}$ y $\varphi \in \mathsf{L}_{\mathsf{Kh}_i}$ por COND. Como Agt es finito y no vacío, esto implica que Act^Γ es numerable, y por lo tanto un conjunto de acciones adecuado para construir un modelo. Con esto, la estructura $\mathfrak{M}^\Gamma = \langle \mathrm{S}^\Gamma, \mathrm{R}^\Gamma, \{\mathrm{U}^\Gamma(i)\}_{i \in \mathsf{Agt}}, \mathrm{V}^\Gamma\rangle$ sobre Act^Γ, Agt y Prop se define como:

- $\mathrm{S}^\Gamma := \{\Delta \in \mathbf{\Phi} \mid \Delta|_{\mathsf{Kh}} = \Gamma|_{\mathsf{Kh}}\}$;
- $\mathrm{R}^\Gamma_{\langle\psi,\varphi\rangle} := \bigcup_{i \in \mathsf{Agt}} \mathrm{R}^\Gamma_{\langle\psi,\varphi\rangle^i}$, con
$$\mathrm{R}^\Gamma_{\langle\psi,\varphi\rangle^i} := \{(\Delta_1,\Delta_2) \in \mathrm{S}^\Gamma \times \mathrm{S}^\Gamma \mid \mathsf{Kh}_i(\psi,\varphi) \in \Gamma, \psi \in \Delta_1, \varphi \in \Delta_2\};$$
- $\mathrm{U}^\Gamma(i) := \Big\{\{\langle\psi,\varphi\rangle\} \mid \langle\psi,\varphi\rangle \in \mathsf{Act}_i^\Gamma\Big\}$; y
- $\mathrm{V}^\Gamma(\Delta) := \{p \in \mathsf{Prop} \mid p \in \Delta\}$.

Como $\Gamma \in \mathbf{\Phi}$, la estructura $\mathfrak{M}^\Gamma$ es del tipo que se requiere.

Proposición 5.2. *La estructura* $\mathfrak{M}^\Gamma = \langle \mathrm{S}^\Gamma, \mathrm{R}^\Gamma, \{\mathrm{U}^\Gamma(i)\}_{i \in \mathsf{Agt}}, \mathrm{V}^\Gamma\rangle$ *es un* LTS^U.

Demostración. Basta con demostrar que cada $\mathrm{U}^\Gamma(i)$ define una partición sobre un subconjunto no vacío de de $2^{(\mathsf{Act}^*)}$. Usando COND, $\mathsf{Kh}_i(\bot,\bot) \in \Gamma$. Por lo tanto, $\langle\bot,\bot\rangle \in \mathsf{Act}_i^\Gamma$ y $\{\langle\bot,\bot\rangle\} \in \mathrm{U}^\Gamma(i)$. Con esto, $\bigcup_{\pi \in \mathrm{U}^\Gamma(i)} \pi \neq \emptyset$. Luego, $\mathrm{U}^\Gamma(i)$ define una partición sobre $\bigcup_{\pi \in \mathrm{U}^\Gamma(i)} \pi$: cada uno de sus conjuntos son mutualmente disjuntos (dado que tienen un solo elemento), y $\emptyset \notin \mathrm{U}^\Gamma(i)$. □

Sea $\Gamma \in \mathbf{\Phi}$; las siguientes propiedades de $\mathfrak{M}^\Gamma$ serán útiles (las pruebas son similares a las dadas en [118]).

Proposición 5.3. *Para todo $\Delta_1, \Delta_2 \in \mathrm{S}^\Gamma$ se tiene que $\Delta_1|_{\mathsf{Kh}} = \Delta_2|_{\mathsf{Kh}}$.*

Demostración. Directo de la definición de S^Γ. □

Proposición 5.4. *Sea $\Delta \in \mathrm{S}^\Gamma$. Si Δ tiene un sucesor con $\mathrm{R}^\Gamma_{\langle\psi,\varphi\rangle}$, entonces cada $\Delta' \in \mathrm{S}^\Gamma$ con $\varphi \in \Delta'$ puede ser alcanzado con $\mathrm{R}^\Gamma_{\langle\psi,\varphi\rangle}$ desde Δ.*

Demostración. Si Δ tiene un sucesor con $\mathrm{R}^\Gamma_{\langle\psi,\varphi\rangle}$, entonces tiene un sucesor con $\mathrm{R}^\Gamma_{\langle\psi,\varphi\rangle^i}$ para algún $i \in \mathsf{Agt}$. Por lo tanto, $\psi \in \Delta$ y $\mathsf{Kh}_i(\psi,\varphi) \in \Gamma$. Luego, todo $\Delta' \in \mathrm{S}^\Gamma$ con $\varphi \in \Delta'$ es tal que $(\Delta, \Delta') \in \mathrm{R}^\Gamma_{\langle\psi,\varphi\rangle^i}$, y por ende tal que $(\Delta, \Delta') \in \mathrm{R}^\Gamma_{\langle\psi,\varphi\rangle}$. □

Proposición 5.5. *Sea φ una fórmula de $\mathsf{L}_{\mathsf{Kh}_i}$. Si $\varphi \in \Delta$ para todo $\Delta \in \mathrm{S}^\Gamma$, entonces $\mathsf{A}\varphi \in \Delta$ para todo $\Delta \in \mathrm{S}^\Gamma$.*

Demostración. Sea $\Delta \in \mathrm{S}^\Gamma \subseteq \mathbf{\Phi}$. Por definición, $\Delta|_{\mathsf{Kh}} \cup \Delta|_{\neg\mathsf{Kh}}$ es un subconjunto de Δ, y por lo tanto es consistente. Más aún todo conjunto que sea extensión maximal consistente $\Delta|_{\mathsf{Kh}} \cup \Delta|_{\neg\mathsf{Kh}}$, Δ', debe cumplir $\Delta|_{\mathsf{Kh}} = \Delta'|_{\mathsf{Kh}}$. Para $(\subseteq)$, $\mathsf{Kh}_i(\psi,\varphi) \in \Delta|_{\mathsf{Kh}}$ implica que $\mathsf{Kh}_i(\psi,\varphi) \in (\Delta|_{\mathsf{Kh}} \cup \Delta|_{\neg\mathsf{Kh}})$, y por lo tanto $\mathsf{Kh}_i(\psi,\varphi) \in \Delta'$, es decir, $\mathsf{Kh}_i(\psi,\varphi) \in \Delta'|_{\mathsf{Kh}}$. Para $(\supseteq)$, se usa la contrapositiva. Si $\mathsf{Kh}_i(\psi,\varphi) \notin \Delta|_{\mathsf{Kh}}$ entonces $\mathsf{Kh}_i(\psi,\varphi) \notin \Delta$, así que $\neg\mathsf{Kh}_i(\psi,\varphi) \in \Delta$, ya que Δ es un conjunto maximal consistente. Con esto, $\neg\mathsf{Kh}_i(\psi,\varphi) \in (\Delta|_{\mathsf{Kh}} \cup \Delta|_{\neg\mathsf{Kh}})$ y $\neg\mathsf{Kh}_i(\psi,\varphi) \in \Delta'$. Luego, $\mathsf{Kh}_i(\psi,\varphi) \notin \Delta'$, dado que Δ' es consistente, y $\mathsf{Kh}_i(\psi,\varphi) \notin \Delta'|_{\mathsf{Kh}}$.

Para demostrar la proposición, supongamos que $\varphi \in \Delta$ para todo $\Delta \in \mathrm{S}^\Gamma$. Sea $\Delta \in \mathrm{S}^\Gamma$, notar que $\Delta|_{\mathsf{Kh}} = \Gamma|_{\mathsf{Kh}}$. Entonces, el conjunto $\Delta|_{\mathsf{Kh}} \cup \Delta|_{\neg\mathsf{Kh}} \cup \{\neg\varphi\}$ es inconsistente. De otra forma se podría extender a un conjunto maximal consistente $\Delta' \in \mathbf{\Phi}$. Por el resultado del anterior párrafo, esto implica que $\Delta'|_{\mathsf{Kh}} = \Delta|_{\mathsf{Kh}}$, $\Delta'|_{\mathsf{Kh}} = \Gamma|_{\mathsf{Kh}}$ y con esto $\Delta' \in \mathrm{S}^\Gamma$. Pero luego, por hipótesis, $\varphi \in \Delta'$, y por construcción, $\neg\varphi \in \Delta'$. Con esto, Δ' inconsistente, lo cual es una contradicción. Por lo tanto, $\Delta|_{\mathsf{Kh}} \cup \Delta|_{\neg\mathsf{Kh}} \cup \{\neg\varphi\}$ es inconsistente y existen conjuntos:

- $\{\mathsf{Kh}_{b_1}(\psi_1,\varphi_1), \ldots, \mathsf{Kh}_{b_n}(\psi_n,\varphi_n)\} \subseteq \Delta|_{\mathsf{Kh}}$ y
- $\{\neg\mathsf{Kh}_{b'_1}(\psi'_1,\varphi'_1), \ldots, \neg\mathsf{Kh}_{b'_m}(\psi'_m,\varphi'_m)\} \subseteq \Delta|_{\neg\mathsf{Kh}}$ tales que

$$\vdash \left(\bigwedge_{k=1}^{n} \mathsf{Kh}_{b_k}(\psi_k,\varphi_k) \wedge \bigwedge_{k=1}^{m} \neg\mathsf{Kh}_{b'_k}(\psi'_k,\varphi'_k)\right) \to \varphi.$$

Usando NECA,

$$\vdash \mathsf{A}\left(\left(\bigwedge_{k=1}^{n} \mathsf{Kh}_{b_k}(\psi_k,\varphi_k) \wedge \bigwedge_{k=1}^{m} \neg\mathsf{Kh}_{b'_k}(\psi'_k,\varphi'_k)\right) \to \varphi\right)$$

y luego, con DISTA y MP,

$$\vdash \mathsf{A}\left(\bigwedge_{k=1}^{n} \mathsf{Kh}_{b_k}(\psi_k,\varphi_k) \wedge \bigwedge_{k=1}^{m} \neg\mathsf{Kh}_{b'_k}(\psi'_k,\varphi'_k)\right) \to \mathsf{A}\varphi.$$

Ahora bien, $\mathsf{Kh}_{b_k}(\psi_k,\varphi_k) \in \Delta|_{\mathsf{Kh}}$ implica ($\mathsf{4KhA}$ y MP) que $\mathsf{AKh}_{b_k}(\psi_k,\varphi_k) \in \Delta$ (para cada $k \in [1 \ldots n]$). De manera análoga, $\neg\mathsf{Kh}_{b'_k}(\psi'_k,\varphi'_k) \in \Delta|_{\mathsf{Kh}}$ implica ($\mathsf{5KhA}$ y MP) que $\mathsf{A}\neg\mathsf{Kh}_{b'_k}(\psi'_k,\varphi'_k) \in \Delta$ (para cada $k \in [1 \ldots m]$). Por lo tanto,

$$\bigwedge_{k=1}^{n} \mathsf{AKh}_{b_k}(\psi_k,\varphi_k) \in \Delta \qquad \text{y} \qquad \bigwedge_{k=1}^{m} \mathsf{A}\neg\mathsf{Kh}_{b'_k}(\psi'_k,\varphi'_k) \in \Delta,$$

se tiene

$$\bigwedge_{k=1}^{n} \mathsf{AKh}_{b_k}(\psi_k,\varphi_k) \wedge \bigwedge_{k=1}^{m} \mathsf{A}\neg\mathsf{Kh}_{b'_k}(\psi'_k,\varphi'_k) \in \Delta, \text{ y } \mathsf{A}\left(\bigwedge_{k=1}^{n} \mathsf{Kh}_{b_k}(\psi_k,\varphi_k) \wedge \bigwedge_{k=1}^{m} \neg\mathsf{Kh}_{b'_k}(\psi'_k,\varphi'_k)\right) \in \Delta.$$

Usando MP, $\mathsf{A}\varphi \in \Delta$. □

Proposición 5.6. *Sean ψ, ψ', φ' fórmulas de $\mathsf{L}_{\mathsf{Kh}_i}$. Si para todo $\Delta \in \mathrm{S}^\Gamma$ tal que $\psi \in \Delta$, tiene un sucesor vía $\mathrm{R}^\Gamma_{\langle\psi',\varphi'\rangle}$. Entonces $\mathsf{A}(\psi \to \psi') \in \Delta$ para todo $\Delta \in \mathrm{S}^\Gamma$.*

Demostración. Sea $\Delta \in \mathrm{S}^\Gamma$. Si $\psi \in \Delta$ entonces, por hipótesis, $(\Delta, \Delta') \in \mathrm{R}^\Gamma_{\langle \psi', \varphi' \rangle}$ para algún Δ'. Luego, por la definición de $\mathrm{R}^\Gamma_{\langle \psi', \varphi' \rangle}$, $\psi' \in \Delta$ y por consistencia maximal $\psi \to \psi' \in \Delta$. Si $\psi \notin \Delta$ entonces $\neg\psi \in \Delta$ por consistencia maximal y por tanto $\psi \to \psi' \in \Delta$. Con esto $\psi \to \psi' \in \Delta$ para todo $\Delta \in \mathrm{S}^\Gamma$. Usando la Proposición 5.5, $\mathsf{A}(\psi \to \psi') \in \Delta$ para todo $\Delta \in \mathrm{S}^\Gamma$. □

Proposición 5.7. *Sea* $\Delta \in \mathrm{S}^\Gamma$ *tal que* $\{\psi, \mathsf{Kh}_i(\psi, \varphi)\} \subset \Delta$*, entonces existe* $\Delta' \in \mathrm{S}^\Gamma$ *tal que* $\varphi \in \Delta'$.

Demostración. Sea $\Delta \in \mathrm{S}^\Gamma$ tal que $\{\psi, \mathsf{Kh}_i(\psi, \varphi)\} \subset \Delta$. Por la contrarrecíproca de TA, se tiene que $\mathsf{E}\psi \in \Delta$. Supongamos que no existe $\Delta' \in \mathrm{S}^\Gamma$ tal que $\varphi \in \Delta'$. Luego, por consistencia maximal, para todo $\Delta' \in \mathrm{S}^\Gamma$, $\neg\varphi \in \Delta'$. Por la Proposición 5.5, para todo $\Delta' \in \mathrm{S}^\Gamma$, $\mathsf{A}\neg\varphi \in \Delta'$ y $\mathsf{A}\neg\varphi \in \Delta$. Como $\{\mathsf{E}\psi, \mathsf{X}(\psi, \varphi), \mathsf{A}\neg\varphi\} \subset \Delta$, por la Proposición 5.1 y consistencia maximal, $\bot \in \Delta$. Pero es contradictorio pues Δ es un conjunto consistente. Por lo tanto, existe $\Delta' \in \mathrm{S}^\Gamma$ tal que $\varphi \in \Delta'$. □

Con estas propiedades se puede demostrar el "truth lemma" para $\mathfrak{M}^\Gamma$. Este lema termina de vincular la semántica de los modelos LTS^U con la deducción sintáctica del sistema $\mathcal{L}^{\mathrm{LTS}^U}_{\mathsf{Kh}_i}$.

Lema 5.1. *Sean* $\Gamma \in \mathbf{\Phi}$, $\mathfrak{M}^\Gamma = \langle \mathrm{S}^\Gamma, \mathrm{R}^\Gamma, \mathrm{U}^\Gamma, \mathrm{V}^\Gamma \rangle$. *Para todo* $\Theta \in \mathrm{S}^\Gamma$ *y* φ *fórmula de* $\mathsf{L}_{\mathsf{Kh}_i}$,

$$\mathfrak{M}^\Gamma, \Theta \models \varphi \qquad \textit{si y sólo si} \qquad \varphi \in \Theta.$$

Demostración. La prueba es por inducción en φ. Los casos atómicos y para los operadores booleanos son directos, por lo que sólo se verá el caso con el operador Kh_i.

- Caso $\mathsf{Kh}_i(\psi, \varphi)$: ($\Rightarrow$) Supongamos que $\mathfrak{M}^\Gamma, \Theta \models \mathsf{Kh}_i(\psi, \varphi)$. Luego existe $\pi \in \mathrm{U}^\Gamma(i)$ tal que: $[\![\psi]\!]^{\mathfrak{M}^\Gamma} \subseteq \mathrm{SE}(\pi)$ y $\mathrm{R}_\pi([\![\psi]\!]^{\mathfrak{M}^\Gamma}) \subseteq [\![\varphi]\!]^{\mathfrak{M}^\Gamma}$. Consideramos dos situaciones:

  - $[\![\psi]\!]^{\mathfrak{M}^\Gamma} = \emptyset$: Por hipótesis inductiva, $\neg\psi \in \Delta$ para todo $\Delta \in \mathrm{S}^\Gamma$. Por la Proposición 5.5, $\mathsf{A}\neg\psi \in \Delta$ para todo $\Delta \in \mathrm{S}^\Gamma$, y con esto $\mathsf{A}(\psi \to \bot) \in \Delta$. Tomando una instancia de SCOND, $\mathsf{A}\neg\psi \to \mathsf{Kh}_i(\psi, \varphi) \in \Delta$. Luego, por MP, se tiene que $\mathsf{Kh}_i(\psi, \varphi) \in \Delta$ para todo $\Delta \in \mathrm{S}^\Gamma$. Por lo tanto, $\mathsf{Kh}_i(\psi, \varphi) \in \Theta$.
  - $[\![\psi]\!]^{\mathfrak{M}^\Gamma} \neq \emptyset$: Luego, $\pi = \{\langle \psi', \varphi' \rangle\}$ y por definición de U^Γ, $\mathsf{Kh}_i(\psi', \varphi') \in \Gamma$ y $\mathrm{R}^\Gamma_{\langle \psi', \varphi' \rangle}$ está definida. Sea $\Delta \in \mathrm{S}^\Gamma$ tal que $\mathfrak{M}^\Gamma, \Delta \models \psi$. Por hipótesis inductiva, $\psi \in \Delta$. Por lo tanto, para todo $\Delta \in \mathrm{S}^\Gamma$, si $\psi \in \Delta$ entonces:
    - Δ tiene un sucesor vía $\mathrm{R}^\Gamma_{\langle \psi', \varphi' \rangle}$, y
    - para todo $\Delta' \in \mathrm{S}^\Gamma$ tal que $(\Delta, \Delta') \in \mathrm{R}^\Gamma_{\langle \psi', \varphi' \rangle}$, $\varphi \in \Delta'$ (hipótesis inductiva).

    Con esto deducimos tres partes:

    1. Como para todo $\Delta \in \mathrm{S}^\Gamma$ tal que $\psi \in \Delta$, tiene un sucesor vía $\mathrm{R}^\Gamma_{\langle \psi', \varphi' \rangle}$, por la Proposición 5.6, $\mathsf{A}(\psi \to \psi') \in \Delta$ para todo $\Delta \in \mathrm{S}^\Gamma$.
    2. Por la Proposición 5.4, cada $\Delta' \in \mathrm{S}^\Gamma$ tal que $\varphi' \in \Delta'$ puede ser alcanzado vía $\mathrm{R}^\Gamma_{\langle \psi', \varphi' \rangle}$ desde un $\Delta \in \mathrm{S}^\Gamma$ tal que $\psi \in \Delta$ (pues, $[\![\psi]\!]^{\mathfrak{M}^\Gamma} \neq \emptyset$) y por lo tanto $\varphi \in \Delta'$. Por consistencia maximal, para todo $\Delta' \in \mathrm{S}^\Gamma$, $\varphi' \to \varphi \in \Delta'$. Usando la Proposición 5.5, para todo $\Delta' \in \mathrm{S}^\Gamma$, $\mathsf{A}(\varphi' \to \varphi) \in \Delta'$.
    3. Dado $\Delta \in \mathrm{S}^\Gamma$, como $\Delta|_{\mathsf{Kh}} = \Gamma|_{\mathsf{Kh}}$ y $\mathsf{Kh}_i(\psi', \varphi') \in \Gamma$, entonces $\mathsf{Kh}_i(\psi', \varphi') \in \Delta$.

    Finalmente, para todo $\Delta \in \mathrm{S}^\Gamma$, $\{\mathsf{A}(\psi \to \psi'), \mathsf{Kh}_i(\psi', \varphi'), \mathsf{A}(\varphi' \to \varphi)\} \subset \Delta$. Usando KhA y MP, $\mathsf{Kh}_i(\psi, \varphi) \in \Theta$.

  ($\Leftarrow$) Supongamos que $\mathsf{Kh}_i(\psi, \varphi) \in \Theta$. Luego, por la Proposición 5.3, $\mathsf{Kh}_i(\psi, \varphi) \in \Delta$ para todo $\Delta \in \mathrm{S}^\Gamma$. Más aún, $\mathsf{Kh}_i(\psi, \varphi) \in \Gamma$ y $\mathrm{R}^\Gamma_{\langle \psi, \varphi \rangle}$ está definida. Para probar que $\mathfrak{M}^\Gamma, \Theta \models \mathsf{Kh}_i(\psi, \varphi)$, consideramos dos casos:

  - No existe Δ tal que $\psi \in \Delta$: Por hipótesis inductiva, $[\![\psi]\!]^{\mathfrak{M}^\Gamma} = \emptyset$. Usando cualquier $\pi \in \mathrm{U}^\Gamma(i)$, se tiene trivialmente que $\mathfrak{M}^\Gamma, \Theta \models \mathsf{Kh}_i(\psi, \varphi)$.
  - Existe Δ tal que $\psi \in \Delta$: Por la Proposición 5.7, existe Δ' tal que $\varphi \in \Delta'$. Por hipótesis inductiva, $\mathfrak{M}^\Gamma, \Delta \models \psi$ y $\mathfrak{M}^\Gamma, \Delta' \models \varphi$. Como está definido, $\pi = \{\langle \psi, \varphi \rangle\} \in \mathrm{U}^\Gamma$ es fuertemente ejecutable en todos los estados donde se cumple ψ (dado que existe un sucesor Δ' vía $\mathrm{R}^\Gamma_{\langle \psi, \varphi \rangle}$) y alcanza desde estos sólo estados donde se cumple φ vía π (por construcción de $\mathrm{R}^\Gamma_{\langle \psi, \varphi \rangle}$). Por lo tanto, $\mathfrak{M}^\Gamma, \Theta \models \mathsf{Kh}_i(\psi, \varphi)$.

□

Y con esto se puede finalmente presentar el resultado objetivo.

Teorema 5.1. *El sistema axiomático* $\mathcal{L}_{\mathsf{Kh}_i}^{\mathrm{LTS}^U}$ *(Tabla 5.1) es correcto y fuertemente completo para* $\mathsf{L}_{\mathsf{Kh}_i}$ *con respecto a la clase de todos los* LTS^U*s.*

Demostración. Para la correctitud, basta con mostrar que los axiomas del sistema son validos y que sus reglas preservan la validez, que resulta ser directo. Para la completitud fuerte, sea Γ' un conjunto de fórmulas consistente con $\mathcal{L}_{\mathsf{Kh}_i}^{\mathrm{LTS}^U}$. Teniendo en cuenta la Observación 5.2, Γ' puede ser extendido a un conjunto maximal consistente con $\mathcal{L}_{\mathsf{Kh}_i}^{\mathrm{LTS}^U}$ $\Gamma \supseteq \Gamma'$. Por el Lema 5.1, $\mathfrak{M}^\Gamma, \Gamma \models \Gamma'$ y Γ' es satisfacible. Dado que $\mathfrak{M}^\Gamma$ es un LTS^U (Proposición 5.2), esto termina de demostrar la completitud fuerte. □

5.3. Notas finales

Algo importante a notar de la construcción del modelo canónico es que cada conjunto de planes indistinguible para un agente es un conjunto de un solo elemento. Por lo tanto, la lógica también es completa con respecto a esta clase particular de modelos. Por otro lado, los modelos LTS^Us son una representación más general y certera desde un punto de vista conceptual. Se podría, por ejemplo, extender el lenguaje de tal forma que está también reflejado por la lógica (por ejemplo, un plan que puede referirse explícitamente a los planes), o definir modalidades parecidas a anuncios públicos para refinar la relación de indistinguibilidad para cada agente. Un trabajo que se ve con más detalle en los Capítulos 8 y 9 que son una extensión del trabajo realizado en [9].

Capítulo 6

Bisimulaciones y expresividad

La experiencia es un sabio hecho a trompicones.

Campoamor

Como se mencionó en la Sección 3.3, las bisimulaciones son una herramienta importante para entender el poder expresivo de una lógica. En este capítulo se definirá la noción de bisimulaciones para $\mathsf{L_{Kh}}$ sobre LTS^Us que toman inspiración de las vistas en [43, 44] y en la sección Sección 3.3 para $\mathsf{L_{Kh}}$ sobre LTSs. Dichas nociones han sido introducidas en [10]. Esto será relevante en capítulos siguientes (ver el Capítulo 8). Además, se comparará la expresividad de los modelos de $\mathsf{L_{Kh}}$ y $\mathsf{L_{Kh}}$, considerando dos clases especiales de modelos LTS^Us, cuyos resultados se han demostrado en [10].

6.1. Bisimulaciones sobre LTS^U

A continuación comenzaremos por generalizar algunas nociones básicas introducidas en la Sección 3.3.

Definición 6.1. Sean $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \{\mathrm{U}(i)\}_{i \in \mathsf{Agt}}, \mathrm{V}\rangle$ un LTS^U sobre Prop, Act y Agt; $\pi \in 2^{(\mathsf{Act}^*)}$ un conjunto de planes, $U, T \subseteq \mathrm{S}$ conjuntos de estados y $i \in \mathsf{Agt}$ un agente.

- Se escribe $U \overset{\pi}{\Rightarrow} T$ sii $U \subseteq \mathrm{SE}(\pi)$ y $\mathrm{R}_\pi(U) \subseteq T$.
- Se escribe $U \overset{i}{\Rightarrow} T$ sii si existe un plan $\pi \in \mathrm{U}(i)$ tal que $U \overset{\pi}{\Rightarrow} T$.

Además, $U \subseteq \mathrm{S}$ es $\mathsf{L_{Kh_i}}$-definible (respectivamente, proposicionalmente definible) en $\mathfrak{M}$ si y sólo si existe una fórmula en $\mathsf{L_{Kh_i}}$ (una fórmula proposicional) φ tal que $U = [\![\varphi]\!]^{\mathfrak{M}}$.

De la definición anterior, se destacan dos cosas. Por un lado, $\overset{i}{\Rightarrow}$ simplifica la semántica de para el operador de saber cómo: $\mathfrak{M}, w \models \mathsf{Kh}_i(\psi, \varphi)$ si y sólo si $[\![\psi]\!]^{\mathfrak{M}} \overset{i}{\Rightarrow} [\![\varphi]\!]^{\mathfrak{M}}$. Por el otro, bajo la semántica basada en LTS^U, $\mathsf{L_{Kh_i}}$-definibilidad también implica definibilidad proposicional. La demostración es análoga a la Proposición 6.1 para LTSs y se basa en el hecho de que Kh_i actúa globalmente.

Proposición 6.1. *Sea $\mathfrak{M}$ un LTS^U. Para todo $U \subseteq \mathrm{D}_{\mathfrak{M}}$, si U es $\mathsf{L_{Kh_i}}$-definible, entonces es proposicionalmente definible.*

Demostración. Las únicas fórmulas no proposicionales en $\mathsf{L_{Kh_i}}$ tienen la modalidad Kh_i. Pero la semántica de Kh_i es global, por lo tanto para todo LTS^U $\mathfrak{M}$ y toda fórmula $\mathsf{Kh}_i(\psi, \varphi)$, el conjunto $[\![\mathsf{Kh}_i(\psi, \varphi)]\!]^{\mathfrak{M}}$ es $\mathrm{D}_{\mathfrak{M}}$ ó $\emptyset$. Por ende, cualquier fórmula $\mathsf{L_{Kh_i}}$ puede ser tranformada en una proposicional semánticamente equivalente reemplazando las subfórmulas de la forma $\mathsf{Kh}_i(\psi, \varphi)$ por $\top$ ó $\bot$ según el caso. $\square$

Con esto ya se puede definir la noción de bisimulación para LTS^Us. Si bien en la colección de relaciones binarias R no está explícitamente mencionada como en la Definición 3.9, está referida mediante la abstracción de "$\overset{i}{\Rightarrow}$" (Definición 6.1).

Definición 6.2 ($\mathsf{L}_{\mathsf{Kh}_i}$-bisimulación)**.** Sean $\mathfrak{M}$ y $\mathfrak{M}'$ dos LTS^Us, siendo S y S′ sus respectivos dominios. Sea $Z \subseteq \mathrm{S} \times \mathrm{S}'$.

- Para $u \in \mathrm{S}$ y $U \subseteq \mathrm{S}$, se definen
$$Z(u) := \{u' \in \mathrm{S}' \mid uZu'\}, \qquad Z(U) := \textstyle\bigcup_{u \in U} Z(u).$$
- Para $u' \in \mathrm{S}'$ y $U' \subseteq \mathrm{S}'$, se definen
$$Z^{-1}(u') := \{u \in \mathrm{S} \mid uZu'\}; \qquad Z^{-1}(U') := \textstyle\bigcup_{u' \in U'} Z^{-1}(u').$$

Un conjunto no vacío $Z \subseteq \mathrm{S} \times \mathrm{S}'$ es una $\mathsf{L}_{\mathsf{Kh}_i}$-bisimulación entre $\mathfrak{M}$ y $\mathfrak{M}'$ si y sólo si wZw' implica lo siguiente.

- **Atom**: $\mathrm{V}(w) = \mathrm{V}'(w')$.
- Kh_i**-Zig**: para todo conjunto *proposicionalmente* definible $U \subseteq \mathrm{S}$, si $U \overset{i}{\Rightarrow} T$ para algún $T \subseteq \mathrm{S}$, entonces existe $T' \subseteq \mathrm{S}'$ tal que

  (B1) $Z(U) \overset{i}{\Rightarrow} T'$, ***(B2)*** $T' \subseteq Z(T)$.
- Kh_i**-Zag**: para todo conjunto *proposicionalmente* definible $U' \subseteq \mathrm{S}'$, si $U' \overset{i}{\Rightarrow} T'$ para algún $T' \subseteq \mathrm{S}'$, entonces existe $T \subseteq \mathrm{S}$ tal que

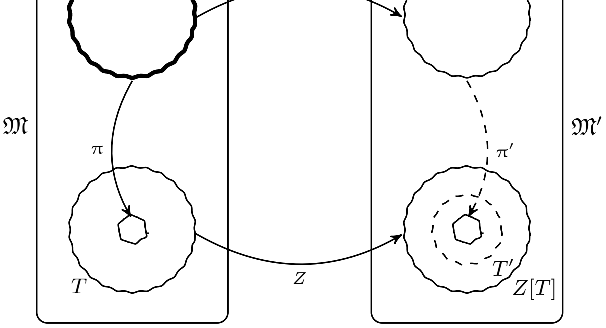
  (B1) $Z^{-1}(U') \overset{i}{\Rightarrow} T$, ***(B2)*** $T \subseteq Z^{-1}(T')$.
- A**-Zig**: para todo $u \in \mathrm{S}$ existe $u' \in \mathrm{S}'$ tal que uZu'.
- A**-Zag**: para todo $u' \in \mathrm{S}'$ existe $u \in \mathrm{S}$ tal que uZu'.

Se escribe $\mathfrak{M}, w \leftrightarroweq_{\mathsf{L}_{\mathsf{Kh}_i}} \mathfrak{M}', w'$ cuando existe una $\mathsf{L}_{\mathsf{Kh}_i}$-bisimulación Z entre $\mathfrak{M}$ y $\mathfrak{M}'$ tal que wZw'.

Análogamente a la Sección 3.3, el siguiente diagrama muestra las condiciones que Kh_i**-Zig** impone. Sea $U \subseteq \mathrm{S}$ un subconjunto proposicionalmente definible, si $U \overset{i}{\Rightarrow} T$ (existe $\pi \in \mathrm{U}(i)$ tal que $U \subseteq \mathrm{SE}(\pi)$ y $\mathrm{R}_\pi(U) \subseteq T$), entonces existe un $T' \subseteq \mathrm{S}'$ tal que

(1) $Z(U) \overset{i}{\Rightarrow} T'$, es decir, para algún $\pi' \in \mathrm{U}'(i)$, T' contiene los estados que son alcanzables por algún estado en $Z(U)$ vía π' ($\mathrm{R}_{\pi'}(Z(U)) \subseteq T'$), y π' es fuertemente ejecutable para todos los estados en $Z(U)$ ($Z(U) \subseteq \mathrm{SE}(\pi')$); y

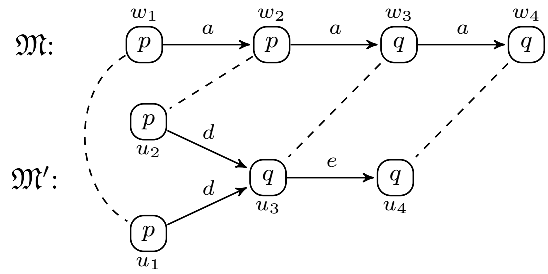
(2) cada estado en T' es la imagen de algún estado en T vía Z ($T' \subseteq Z(T)$).



Las condiciones de Kh_i-**Zag** funcionan en la otra dirección: para cualquier conjunto proposicionalmente definible $U' \subseteq \mathrm{S}'$, que cumpla $U' \overset{i}{\Rrightarrow} T'$ para algún $T' \subseteq \mathrm{S}'$, entonces existe $T \subseteq \mathrm{S}$ tal que ***(1)*** $Z^{-1}(U') \overset{i}{\Rrightarrow} T$ y ***(2)*** $T \subseteq Z^{-1}(T')$. El diagrama asociado es análogo al anterior.

Además de esto, las dos condiciones en Kh_i-**Zig** pueden simplificarse como una única condición: $Z(U) \overset{i}{\Rrightarrow} Z(T)$. La presentación anterior se parece más a la definición estándar de una bisimulación: si U tiene un "sucesor mediante i" T, entonces su imagen U' vía Z también tiene un "sucesor mediante i" llamado T' ($Z(U) \overset{i}{\Rrightarrow} T'$), y estos sucesores "coinciden bisimilarmente" ($T' \subseteq Z(T)$). El caso de Kh_i-**Zag** es análogo.

Ejemplo 6.1. Sean $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \{\mathrm{U}(i)\}_{i\in\mathsf{Agt}}, \mathrm{V}\rangle$ y $\mathfrak{M}' = \langle \mathrm{S}', \mathrm{R}', \{\mathrm{U}'(i)\}_{i\in\mathsf{Agt}}, \mathrm{V}'\rangle$ los siguientes LTS^Us con $\mathrm{U}(i) = \{\{aa\}\}$ y $\mathrm{U}'(i) = \{\{d, de\}\}$:



Tomando $Z = \{(w_1, u_1), (w_2, u_2), (w_3, u_3), (w_4, u_4)\}$, representado en el gráfico con las líneas punteadas, se puede observar que Z cumple **Atom**, A-**Zig** y A-**Zag**. Más aún, Kh_i-**Zig** también. Todo subconjunto $U \subseteq \mathrm{S}$ proposicionalmente definible satisface las condiciones como se muestra en la siguiente tabla, definida similarmente a la vista en el Ejemplo 3.11.

U	Def. como	T	$U \overset{i}{\Rrightarrow} T$ por	$Z(U)$	$Z(U) \overset{i}{\Rrightarrow} Z(T)$ por	$Z(T)$
$\emptyset$	$\bot$	$\emptyset$	$\{aa\}$	$\emptyset$	$\{d, de\}$	$\emptyset$
$\{w_1, w_2\}$	p	$\{w_3, w_4\}$	$\{aa\}$	$\{u_1, u_2\}$	$\{d, de\}$	$\{u_3, u_4\}$

En la misma, se listan exhaustivamente todos los conjuntos U y T en el modelo, y sus correspondientes imágenes $Z(U)$ y $Z(T)$ tales que $U \overset{i}{\Rrightarrow} T$. La segunda columna especifica qué fórmula proposicional define el conjunto U mientras que la cuarta y sexta columnas enumeran los testigos para cada relación $U \overset{i}{\Rrightarrow} T$ y $Z(U) \overset{i}{\Rrightarrow} Z(T)$ respectivamente. Con un razonamiento análogo, se tiene que Kh_i-**Zag** se cumple y Z es una $\mathsf{L}_{\mathsf{Kh}_i}$-bisimulación entre $\mathfrak{M}$ y $\mathfrak{M}'$. Además, $\mathfrak{M}, w_1 \underline{\leftrightarrow}_{\mathsf{L}_{\mathsf{Kh}_i}} \mathfrak{M}', u_1$ y $\mathfrak{M}, w_3 \underline{\leftrightarrow}_{\mathsf{L}_{\mathsf{Kh}_i}} \mathfrak{M}', u_3$ dado que $w_1 Z u_1$ y $w_3 Z u_3$ respectivamente. $\square$

Así como sucedió en la Sección 3.3, se define la noción de equivalencia de modelos con respecto a $\mathsf{L}_{\mathsf{Kh}_i}$.

Definición 6.3 ($\mathsf{L}_{\mathsf{Kh}_i}$-equivalencia)**.** Sean $\mathfrak{M}$ y $\mathfrak{M}$ dos $\mathrm{LTS}^U s$, $w \in \mathrm{D}_{\mathfrak{M}}$ y $w' \in \mathrm{D}_{\mathfrak{M}'}$. $\mathfrak{M}, w$ y $\mathfrak{M}', w'$ son $\mathsf{L}_{\mathsf{Kh}_i}$-*equivalentes* ($\mathfrak{M}, w \leftrightsquigarrow_{\mathsf{L}_{\mathsf{Kh}_i}} \mathfrak{M}', w'$) si y sólo si, para cada $\varphi \in \mathsf{L}_{\mathsf{Kh}_i}$,

$$\mathfrak{M}, w \models \varphi \quad \text{sii} \quad \mathfrak{M}', w' \models \varphi.$$

Una vez introducidas estas nociones, podemos establecer los teoremas de correspondencia para $\mathsf{L}_{\mathsf{Kh}_i}$.

Teorema 6.1 (Invarianza bajo bisimulaciones)**.** *Sean $\mathfrak{M}$ y $\mathfrak{M}$ $\mathrm{LTS}^U s$, $w \in \mathrm{D}_{\mathfrak{M}}$ y $w' \in \mathrm{D}_{\mathfrak{M}'}$. Entonces,*

$$\mathfrak{M}, w \underline{\leftrightarrow}_{\mathsf{L}_{\mathsf{Kh}_i}} \mathfrak{M}', w' \quad \textit{implica} \quad \mathfrak{M}, w \leftrightsquigarrow_{\mathsf{L}_{\mathsf{Kh}_i}} \mathfrak{M}', w'.$$

Demostración. Sean $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \{\mathrm{U}(i)\}_{i\in\mathsf{Agt}}, \mathrm{V}\rangle$ y $\mathfrak{M}' = \langle \mathrm{S}', \mathrm{R}', \{\mathrm{U}'(i)\}_{i\in\mathsf{Agt}}, \mathrm{V}'\rangle$. Sea $Z \subseteq (\mathrm{S}\times\mathrm{S}')$ la $\mathsf{L}_{\mathsf{Kh}_i}$-bisimulación entre $\mathfrak{M}, w$ y $\mathfrak{M}', w'$, tal que wZw'. La prueba de $\mathsf{L}_{\mathsf{Kh}_i}$-equivalencia es por inducción en la estructura de las fórmulas de $\mathsf{L}_{\mathsf{Kh}_i}$. Los casos para las proposiciones atómicas y operadores booleanos son estándar. Por lo cual, sólo quedan las fórmulas de la forma $\mathsf{Kh}_i(\psi, \varphi)$.

Para este caso la hipótesis inductiva establece que para $u \in \mathrm{S}$, $u' \in \mathrm{S}'$ y χ subfórmula propia de $\mathsf{Kh}_i(\psi, \varphi)$, si uZu' entonces $u \in [\![\chi]\!]^{\mathfrak{M}}$ sii $u' \in [\![\chi]\!]^{\mathfrak{M}'}$.

Sea $w \in [\![\mathsf{Kh}_i(\psi, \varphi)]\!]^{\mathfrak{M}}$, por la semántica y la Definición 6.1, existe un conjunto de planes $\pi \in \mathrm{U}(i)$ tal que $[\![\psi]\!]^{\mathfrak{M}} \stackrel{\pi}{\Rightarrow} [\![\varphi]\!]^{\mathfrak{M}}$. Con esto, $[\![\psi]\!]^{\mathfrak{M}} \stackrel{i}{\Rightarrow} [\![\varphi]\!]^{\mathfrak{M}}$. Algo importante para tener en cuenta es que $Z([\![\chi]\!]^{\mathfrak{M}}) = [\![\chi]\!]^{\mathfrak{M}'}$ se cumple para $\chi \in \{\psi, \varphi\}$.

($\subseteq$) Si $v' \in Z([\![\chi]\!]^{\mathfrak{M}})$, entonces existe $v \in [\![\chi]\!]^{\mathfrak{M}}$ tal que vZv'. Por hipótesis inductiva $v' \in [\![\chi]\!]^{\mathfrak{M}'}$.

($\supseteq$) Si $v' \in [\![\chi]\!]^{\mathfrak{M}'}$, entonces por A-**Zag** existe v tal que vZv'. Por hipótesis inductiva $v \in [\![\chi]\!]^{\mathfrak{M}}$ y con esto, $v' \in Z([\![\chi]\!]^{\mathfrak{M}})$.

Dado que $[\![\psi]\!]^{\mathfrak{M}} \stackrel{i}{\Rightarrow} [\![\varphi]\!]^{\mathfrak{M}}$. El conjunto $[\![\psi]\!]^{\mathfrak{M}}$ es obviamente $\mathsf{L}_{\mathsf{Kh}_i}$-definible, y por la Proposición 6.1 es proposicionalmente definible. Usando la condición Kh_i-**Zig**, existe $T' \subseteq \mathrm{S}'$ tal que ***(B1)*** $Z([\![\psi]\!]^{\mathfrak{M}}) \stackrel{i}{\Rightarrow} T'$ y ***(B2)*** $T' \subseteq Z([\![\varphi]\!]^{\mathfrak{M}})$. Por lo tanto, $Z([\![\psi]\!]^{\mathfrak{M}}) \stackrel{i}{\Rightarrow} Z([\![\varphi]\!]^{\mathfrak{M}})$ y usando el resultado anterior, $[\![\psi]\!]^{\mathfrak{M}'} \stackrel{i}{\Rightarrow} [\![\varphi]\!]^{\mathfrak{M}'}$. Con esto, $w' \in [\![\mathsf{Kh}_i(\psi, \varphi)]\!]^{\mathfrak{M}'}$.

La dirección de $w' \in [\![\mathsf{Kh}_i(\psi, \varphi)]\!]^{\mathfrak{M}'}$ a $w \in [\![\mathsf{Kh}_i(\psi, \varphi)]\!]^{\mathfrak{M}}$ sigue un argumento similar, usando A-**Zig** y Kh_i-**Zag**. □

Ejemplo 6.2. Tomando el Ejemplo 6.1, $\mathfrak{M}, w_1 \leftrightsquigarrow_{\mathsf{L}_{\mathsf{Kh}_i}} \mathfrak{M}', u_1$ y $\mathfrak{M}, w_3 \leftrightsquigarrow_{\mathsf{L}_{\mathsf{Kh}_i}} \mathfrak{M}', u_3$. □

Como sucedía en la Sección 3.3, la vuelta de la implicación del Teorema 6.1 no se cumple para modelos arbitrarios (ver [44, Sección 2] y el Ejemplo 3.13 con un único agente i tal que $\mathrm{U}(i) = \mathrm{U}'(i) = \{\{\epsilon\}\}$). Para que esto suceda, se deberá restringir a una clase de modelos particular, generalmente conocidas como clases Hennessy-Milner. Una vez más, debido a la presencia de la modalidad universal A, la clase de los modelos finitos (en el sentido de que tienen dominios finitos) resulta ser una buena candidata.

Definición 6.4. Se define la clase de modelos LTS^Us finitos en dominio (**FD**):

$$\mathbf{M_{FD}} := \{\mathfrak{M} \mid \mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \mathrm{U}, \mathrm{V} \rangle \text{ es un } \mathrm{LTS}^U \text{ tal que S es finito}\}.$$

Un detalle importante es que no hay restricciones sobre la relación de incertidumbre de los agentes (los conjuntos $\mathrm{U}(i)$). Para terminar esta sección, el siguiente teorema demuestra que $\mathbf{M_{FD}}$ es una clase Hennessy-Milner.

Teorema 6.2 ($\mathsf{L}_{\mathsf{Kh}_i}$-equivalencia implica $\mathsf{L}_{\mathsf{Kh}_i}$-bisimilaridad)**.** *Sean $\mathfrak{M}, \mathfrak{M}' \in \mathbf{M_{FD}}$, $w \in \mathrm{D}_{\mathfrak{M}}$ y $w' \in \mathrm{D}_{\mathfrak{M}'}$. Se tiene que*

$$\mathfrak{M}, w \leftrightsquigarrow_{\mathsf{L}_{\mathsf{Kh}_i}} \mathfrak{M}', w' \quad \textit{implica} \quad \mathfrak{M}, w \,\underline{\leftrightarrow}_{\mathsf{L}_{\mathsf{Kh}_i}} \mathfrak{M}', w'.$$

Demostración. Sean $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \{\mathrm{U}(i)\}_{i \in \mathsf{Agt}}, \mathrm{V} \rangle$ y $\mathfrak{M}' = \langle \mathrm{S}', \mathrm{R}', \{\mathrm{U}'(i)\}_{i \in \mathsf{Agt}}, \mathrm{V}' \rangle$. La estrategia es mostrar que la relación $\leftrightsquigarrow_{\mathsf{L}_{\mathsf{Kh}_i}}$ es de por sí una $\mathsf{L}_{\mathsf{Kh}_i}$-bisimulación. Con esto se define

$$Z := \{(v, v') \in (\mathrm{S} \times \mathrm{S}') \mid \mathfrak{M}, v \leftrightsquigarrow_{\mathsf{L}_{\mathsf{Kh}_i}} \mathfrak{M}', v'\}.$$

Luego wZw' implica que w y w' cumplen las mismas fórmulas en $\mathsf{L}_{\mathsf{Kh}_i}$. Para ver que Z cumple las condiciones, sea $(w, w') \in Z$.

- **Atom**. Dado que w y w' cumplen las mismas fórmulas de $\mathsf{L}_{\mathsf{Kh}_i}$, también cumplen todas las variables proposicionales.
- A-**Zig**. Sea $v \in \mathrm{S}$ supongamos que no existe $v' \in \mathrm{S}'$ tal que vZv'. De la definición de Z, para cada $v'_i \in \mathrm{S}' = \{v'_1, \ldots, v'_n\}$ (dado que $\mathfrak{M}'$ es finito) existe una fórmula en $\mathsf{L}_{\mathsf{Kh}_i}$ θ_i tal que $\mathfrak{M}, v \models \theta_i$ pero $\mathfrak{M}', v'_i \not\models \theta_i$. Sea $\theta := \theta_1 \wedge \cdots \wedge \theta_n$. Claramente, $\mathfrak{M}, v \models \theta$. Sin embargo, $\mathfrak{M}', v'_i \not\models \theta$ para cada $v'_i \in \mathrm{S}'$, dado que cada uno hace falso a su propio θ_i. Luego, $\mathfrak{M}, w \models \mathsf{E}\theta$ pero $\mathfrak{M}', w' \not\models \mathsf{E}\theta$, contradiciendo wZw'.
- A-**Zag**. Análogo al caso A-**Zig**.
- Kh_i-**Zig**. Sea $[\![\psi]\!]^{\mathfrak{M}} \subseteq \mathrm{S}$ un conjunto proposicionalmente definible (ψ es proposicional), supongamos que $[\![\psi]\!]^{\mathfrak{M}} \stackrel{i}{\Rightarrow} T$ para algún $T \subseteq \mathrm{S}$. Se debe demostrar que existe $T' \subseteq \mathrm{S}'$ tal que

(B1) $Z([\![\psi]\!]^{\mathfrak{M}}) \overset{i}{\Rightarrow} T'$, ***(B2)*** $T' \subseteq Z(T)$.

Se puede ver que $Z([\![\psi]\!]^{\mathfrak{M}}) = [\![\psi]\!]^{\mathfrak{M}'}$. Para ($\supseteq$), sea $u' \in [\![\psi]\!]^{\mathfrak{M}'}$. Usando A-**Zag** (demostrado más arriba), existe $u \in \mathrm{S}$ tal que uZu'. De la definición de Z, $u \in [\![\psi]\!]^{\mathfrak{M}}$ y con esto $u' \in Z([\![\psi]\!]^{\mathfrak{M}})$. Para ($\subseteq$), sea $u' \in Z([\![\psi]\!]^{\mathfrak{M}})$. Entonces existe $u \in [\![\psi]\!]^{\mathfrak{M}}$ tal que uZu', y por lo tanto, usando la definición de Z, $u' \in [\![\psi]\!]^{\mathfrak{M}'}$. Con esto, se debe demostrar que existe $T' \subseteq \mathrm{S}'$ tal que

(B1) $[\![\psi]\!]^{\mathfrak{M}'} \overset{i}{\Rightarrow} T'$, ***(B2)*** $T' \subseteq Z(T)$.

Consideremos dos casos.

1. Si $[\![\psi]\!]^{\mathfrak{M}} = \emptyset$, entonces $[\![\psi]\!]^{\mathfrak{M}'} = Z([\![\psi]\!]^{\mathfrak{M}}) = \emptyset$ y con esto $T' = \emptyset$ es el perfecto candidato ya que cumple

   (B1) $\emptyset \overset{i}{\Rightarrow} \emptyset$ (dado que $\mathrm{U}(i) \neq \emptyset$), ***(B2)*** $\emptyset \subseteq Z(T)$.

2. Si $[\![\psi]\!]^{\mathfrak{M}} \neq \emptyset$, entonces $T \neq \emptyset$ por la definición de $[\![\psi]\!]^{\mathfrak{M}} \overset{i}{\Rightarrow} T$. Supongamos que no existe $T' \subseteq \mathrm{S}'$ tal que cumpla ***(B1)*** y ***(B2)***. En otras palabras, para todo $T' \subseteq \mathrm{S}'$ tal que satisface ***(B1)***, no cumple ***(B2)***. Esto quiere decir que para todo $T' \subseteq \mathrm{S}'$ que cumple $[\![\psi]\!]^{\mathfrak{M}'} \overset{i}{\Rightarrow} T'$ tiene un estado $v'_{T'} \in T'$ tal que para todo $v \in T$, $vZv'_{T'}$ no sucede. Usando la definición de Z, lo último significa que todo estado en T puede ser distinguible de este $v'_{T'}$ por una fórmula $\mathsf{L}_{\mathsf{Kh}_i}$. Por lo tanto, dado cualquier $T' \subseteq \mathrm{S}'$ que cumple $[\![\psi]\!]^{\mathfrak{M}'} \overset{i}{\Rightarrow} T'$, existe un estado $v'_{T'} \in T'$ que para cada $v \in T$, hay una fórmula $\theta^{v}_{v'_{T'}}$ tal que $\mathfrak{M}, v \models \theta^{v}_{v'_{T'}}$ pero $\mathfrak{M}', v'_{T'} \not\models \theta^{v}_{v'_{T'}}$. Luego, para cada $v'_{T'}$ en cada T' se define

   $$\theta_{T'} := \bigvee_{v \in T} \theta^{v}_{v'_{T'}} \qquad \text{y luego} \qquad \theta := \bigwedge_{\{T' \subseteq \mathrm{S}' \mid [\![\psi]\!]^{\mathfrak{M}'} \overset{i}{\Rightarrow} T'\}} \theta_{T'},$$

   Dado que S es finito, $\theta_{T'}$ es una fórmula. Como S' es finito, $\{T' \subseteq \mathrm{S}' \mid [\![\psi]\!]^{\mathfrak{M}'} \overset{i}{\Rightarrow} T'\}$ también lo es y con esto θ_T es una fórmula. Se tiene entonces que $\theta_{T'}$ es la disyunción de fórmulas que distinguen $v'_{T'}$ de cualquier estado en T, y θ es la conjunción de todas estas. Dado que $T \neq \emptyset$, $\theta_{T'}$ no colapsa a $\bot$. Sin embargo θ sí puede colapsar a $\top$ dado que $\{T' \subseteq \mathrm{S}' \mid [\![\psi]\!]^{\mathfrak{M}'} \overset{i}{\Rightarrow} T'\}$ puede ser vacío.

   - Si $\{T' \subseteq \mathrm{S}' \mid [\![\psi]\!]^{\mathfrak{M}'} \overset{i}{\Rightarrow} T'\} = \emptyset$, se considera la fórmula $\mathsf{Kh}_i(\psi, \top)$. Como $[\![\psi]\!]^{\mathfrak{M}} \overset{i}{\Rightarrow} T$ y $T \subseteq \mathrm{S} = [\![\top]\!]^{\mathfrak{M}}$, se tiene que $\mathfrak{M}, w \models \mathsf{Kh}_i(\psi, \top)$. Aún así, $\mathfrak{M}', w' \not\models \mathsf{Kh}_i(\psi, \top)$ ya que no hay $T' \subseteq \mathrm{S}' = [\![\top]\!]^{\mathfrak{M}'}$ tal que $[\![\psi]\!]^{\mathfrak{M}'} \overset{i}{\Rightarrow} T'$. Esto contradice $\mathsf{L}_{\mathsf{Kh}_i}$-equivalencia de w y w'.
   - Si $\{T' \subseteq \mathrm{S}' \mid [\![\psi]\!]^{\mathfrak{M}'} \overset{i}{\Rightarrow} T'\} \neq \emptyset$, entonces θ no colapsa a $\top$. Notar que cada $v \in T$ cumple su propio $\theta^{v}_{v'_{T'}}$ en cada $\theta_{T'}$, y con ello satisface θ. Por ende, $T \subseteq [\![\theta]\!]^{\mathfrak{M}}$. Dado que $[\![\psi]\!]^{\mathfrak{M}} \overset{i}{\Rightarrow} T$ y que las fórmulas Kh_i son globales, se tiene que $\mathfrak{M}, w \models \mathsf{Kh}_i(\psi, \theta)$. Sin embargo, para cada T' en $\{T' \subseteq \mathrm{S}' \mid [\![\psi]\!]^{\mathfrak{M}'} \overset{i}{\Rightarrow} T'\}$, el estado $v'_{T'}$, que no coincide con ningún estado $v \in T$, no cumple ninguno de los $\theta^{v}_{v'_{T'}}$, falsificando $\theta_{T'}$ y por ende también a θ. Con esto, para todo $T' \subseteq \mathrm{S}'$ tal que $[\![\psi]\!]^{\mathfrak{M}'} \overset{i}{\Rightarrow} T'$, $\mathfrak{M}', t' \not\models \theta$ para algún estado $t' \in T'$. Por lo tanto, $[\![\psi]\!]^{\mathfrak{M}'} \overset{i}{\Rightarrow} T' \not\subseteq [\![\theta]\!]^{\mathfrak{M}'}$ y, dado que las fórmulas Kh_i son globales, $\mathfrak{M}', w' \not\models \mathsf{Kh}_i(\psi, \theta)$, contradiciendo la $\mathsf{L}_{\mathsf{Kh}_i}$-equivalencia de w y w'.

- Kh_i-**Zag**. Análogo al caso Kh_i-**Zig**.

$\square$

6.2. Comparación entre $\mathsf{L_{Kh}}$ y $\mathsf{L_{Kh_{\mathit{i}}}}$

Cuando se define una lógica utilizando otra como punto de partida, es necesario comparar la expresividad de ambas ya sea si alguna de estas es capaz de “captar” o “contener” los modelos de la otra o si simplemente resultan ser demasiado distintas entre sí. Notar que, en el caso de $\mathsf{L_{Kh}}$ y $\mathsf{L_{Kh_{\mathit{i}}}}$, los modelos de ambas lógicas son diferentes por lo que una comparación de expresividad relativa carece de sentido. Por ende, en esta sección compararemos si una lógica es capaz de capturar los mismos teoremas y fórmulas válidas que su contraparte o no. Específicamente, se mostrará que $\mathsf{L_{Kh_{\mathit{i}}}}$ es más débil que $\mathsf{L_{Kh}}$. Luego se explorarán dos clases de modelos LTS^Us con los cuales podemos capturar la semántica exacta de la Definición 3.7. Como consecuencia, el sistema axiomático en la Tabla 3.1 es correcto y fuertemente completo con respecto a ambas clases, mostrando así que la lógica basada en LTS^Us generaliza a la basada en LTSs. Para esta comparación, nos restringiremos a los LTS^U con un solo agente, definiendo una sola modalidad Kh sin subíndices. Los LTS^Us a partir de esta sección hasta el final del capítulo serán de la forma $\langle \mathrm{S}, \mathrm{R}, \mathrm{U}, \mathrm{V}\rangle$, con U representando la indistinguibilidad del único agente considerado y no una función. Con esto, la definición de A en $\mathsf{L_{Kh_{\mathit{i}}}}$ termina siendo exactamente como la de $\mathsf{L_{Kh}}$.

Para empezar, mostraremos que el sistema axiomático $\mathcal{L}^{\mathrm{LTS}^U}_{\mathsf{Kh}_i}$ es deducible del sistema $\mathcal{L}^{\mathrm{LTS}}_{\mathsf{Kh}}$. Esto se puede observar como una consecuencia de la siguiente propiedad.

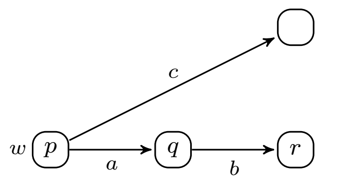
Proposición 6.2. *KhA es un teorema de* $\mathcal{L}^{\mathrm{LTS}}_{\mathsf{Kh}}$.

Demostración. Se tiene que $\vdash (\mathsf{Kh}_i(\chi,\psi) \wedge \mathsf{Kh}_i(\psi,\varphi)) \to \mathsf{Kh}_i(\chi,\varphi)$ y $\vdash (\mathsf{Kh}_i(\chi,\varphi) \wedge \mathsf{Kh}_i(\varphi,\theta)) \to \mathsf{Kh}_i(\chi,\theta)$, dos instancias de COMPKh. Por razonamiento proposicional, $\vdash (\mathsf{Kh}_i(\chi,\psi) \wedge \mathsf{Kh}_i(\psi,\varphi) \wedge \mathsf{Kh}_i(\varphi,\theta)) \to \mathsf{Kh}_i(\chi,\theta)$. Usando dos instancias de EMP $\vdash \mathsf{A}(\chi \to \psi) \to \mathsf{Kh}_i(\chi,\varphi)$ y $\vdash \mathsf{A}(\varphi \to \theta) \to \mathsf{Kh}_i(\varphi,\theta)$, se tiene por razonamiento proposicional que $\vdash (\mathsf{A}(\chi \to \psi) \wedge \mathsf{Kh}_i(\psi,\varphi) \wedge \mathsf{A}(\varphi \to \theta)) \to \mathsf{Kh}_i(\chi,\theta)$. $\square$

Con esto, el operador Kh basado en LTSs es al menos tan fuerte como su contraparte basada en LTS^Us. Dado que todos los axiomas de la Tabla 5.1 son teoremas de la Tabla 3.1, se tiene que toda fórmula válida bajo LTS^Us es también válida bajo LTSs. La siguiente proposición, análoga al Ejemplo 5.1, muestra que la conversa no se cumple.

Proposición 6.3. *Los axiomas EMP y COMPKh no son válidos en la semántica basada en* LTS^U*s.*

Demostración. Consideremos el LTS^U, con la siguiente colección de conjuntos de planes para el agente (el conjunto U).



$$\mathrm{U} = \left\{ \begin{array}{c} \{a\},\ \{b\} \\ \{ab, c\} \end{array} \right\}$$

Para EMP se cumple que $\mathsf{A}(p \to p)$. Sin embargo, $\mathsf{Kh}(p,p)$ no se cumple ya que no hay $\pi \in \mathrm{U}$ que lleve desde estados p a estados p. Generalmente, EMP es válido para LTSs dado que el plan vacío ϵ es fuertemente ejecutable en todos los estados y siempre está disponible como testigo. Sin embargo, en un LTS^U, ϵ puede no estar disponible para el agente (es decir, $\epsilon \notin \mathrm{P}$), y aún estando puede ser indistinguible de otros planes que se comportan de manera diferente.

Con respecto a COMPKh, notar que $\mathsf{Kh}(p,q)$ y $\mathsf{Kh}(q,r)$ se cumplen, siendo $\{a\}$ y $\{b\}$ sus correspondientes testigos. Pero no existe $\pi \in \mathrm{U}$ con sólo planes que al ejecutarse en estados p lleguen únicamente a estados r. Por lo tanto, $\mathsf{Kh}(p,r)$ no se cumple. En general, COMPKh es válida sobre LTSs porque la composición secuencial de planes que hacen verdadera las dos partes de la conjunción en el antecedente es un testigo que hace verdadero el consecuente. Sin embargo, en un LTS^U, esta composición nuevamente puede no estar disponible o ser indistinguible de otros planes que no sean adecuados. $\square$

De estas dos observaciones se tiene que Kh bajo LTS^Us es estrictamente más débil que Kh bajo LTSs. Con esto se asevera que agregar incertidumbre entre planes cambia la lógica.

6.2.1. Una clase simple de LTS^Us

Como mostraremos a continuación, la semántica basada en LTS^U es lo suficientemente general para capturar la semántica para LTSs. Dada la discusión de la Proposición 6.3, hay una clase obvia de LTS^Us en la cual EMP y COMPKh son válidas: la clase de LTS^Us en la cual el agente tiene disponible todos los planes y puede distinguir a todos entre sí, formalizada de la siguiente manera.

Definición 6.5. Se define la clase de modelos LTS^Us sin incertidumbre (**NU**):

$$\mathbf{M_{NU}} := \{\mathfrak{M} \mid \mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \mathrm{U}, \mathrm{V}\rangle \text{ es un } \mathrm{LTS}^U \text{ tal que } \mathrm{U} = \{\{\sigma\} \mid \sigma \in \mathsf{Act}^*\}\}.$$

En efecto, para modelos en $\mathbf{M_{NU}}$, el plan ϵ siempre está disponible y es distinguible de otros planes (cumpliendo EMP) y de $\{\sigma_1\} \in \mathrm{U}$ y $\{\sigma_2\} \in \mathrm{U}$ se sigue que $\{\sigma_1\sigma_2\} \in \mathrm{U}$ (cumpliendo COMPKh). Por lo tanto, la siguiente proposición establece que un agente en un LTS es exactamente un agente en un LTS^U que tiene todos los planes disponibles y no tiene incertidumbre sobre ninguno de estos. Esta clase es suficiente para mostrar cómo la lógica basada en incertidumbre puede capturar la original.

Proposición 6.4. *Las siguientes propiedades se cumplen.*

1. *Dado un modelo* $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \mathrm{U}, \mathrm{V}\rangle$ *en* $\mathbf{M_{NU}}$ *y el* LTS $\mathfrak{L}_{\mathfrak{M}} = \langle \mathrm{S}, \mathrm{R}, \mathrm{V}\rangle$*, se tiene que* $[\![\varphi]\!]^{\mathfrak{M}} = [\![\varphi]\!]^{\mathfrak{L}_{\mathfrak{M}}}$*, para cada* $\varphi \in \mathsf{L}_{\mathsf{Kh}}$.
2. *Dado un* LTS $\mathfrak{L} = \langle \mathrm{S}, \mathrm{R}, \mathrm{V}\rangle$ *y el modelo* $\mathfrak{M}_{\mathfrak{L}} = \langle \mathrm{S}, \mathrm{R}, \mathrm{U}, \mathrm{V}\rangle$ *con* $\mathrm{U} = \{\{\sigma\} \mid \sigma \in \mathsf{Act}^*\}$*, se tiene que* $\mathfrak{M}_{\mathfrak{L}} \in \mathbf{M_{NU}}$*, y es tal que* $[\![\varphi]\!]^{\mathfrak{L}} = [\![\varphi]\!]^{\mathfrak{M}_{\mathfrak{L}}}$*, para cada* $\varphi \in \mathsf{L}_{\mathsf{Kh}}$.

Demostración. En ambos casos, la prueba es por inducción estructural en la fórmula $\varphi \in \mathsf{L}_{\mathsf{Kh}}$.

1. El caso para proposiciones atómicas es directo dado que $\mathfrak{M}$ y $\mathfrak{L}_{\mathfrak{M}}$ tienen la misma valuación atómica. Los casos con los operadores booleanos se cumplen usando la hipótesis inductiva. Para el caso Kh, $(\subseteq)$ sea $w \in [\![\mathsf{Kh}(\psi,\varphi)]\!]^{\mathfrak{M}}$, se tiene que existe $\pi \in \mathrm{U}$ tal que **(Kh-1)** $[\![\psi]\!]^{\mathfrak{M}} \subseteq \mathrm{SE}(\pi)$ y **(Kh-2)** $\mathrm{R}_\pi([\![\psi]\!]^{\mathfrak{M}}) \subseteq [\![\varphi]\!]^{\mathfrak{M}}$. Como $\mathfrak{M} \in \mathbf{M_{NU}}$, necesariamente se tiene que $\pi = \{\sigma\}$ para algún $\sigma \in \mathsf{Act}^*$.

   Por un lado, usando hipótesis inductiva, $[\![\psi]\!]^{\mathfrak{L}_{\mathfrak{M}}} = [\![\psi]\!]^{\mathfrak{M}}$. Luego, por **(Kh-1)**, $[\![\psi]\!]^{\mathfrak{M}} \subseteq \mathrm{SE}(\pi)$. Aplicando la definición de SE para un conjunto de planes, $\mathrm{SE}(\pi) = \mathrm{SE}(\{\sigma\}) = \mathrm{SE}(\sigma)$. Con esto, $[\![\psi]\!]^{\mathfrak{L}_{\mathfrak{M}}} \subseteq \mathrm{SE}(\sigma)$.

   Por otro lado, usando la hipótesis inductiva y la definición de R para conjuntos de planes, $\mathrm{R}_\sigma([\![\psi]\!]^{\mathfrak{L}_{\mathfrak{M}}}) = \mathrm{R}_\pi([\![\psi]\!]^{\mathfrak{M}})$. Aplicando **(Kh-2)** e hipótesis inductiva nuevamente, $\mathrm{R}_\pi([\![\psi]\!]^{\mathfrak{M}}) \subseteq [\![\varphi]\!]^{\mathfrak{M}} = [\![\varphi]\!]^{\mathfrak{L}_{\mathfrak{M}}}$. Con esto, $\mathrm{R}_\sigma([\![\psi]\!]^{\mathfrak{L}_{\mathfrak{M}}}) \subseteq [\![\varphi]\!]^{\mathfrak{L}_{\mathfrak{M}}}$.

   Por lo tanto, **(Kh-1)** $[\![\psi]\!]^{\mathfrak{L}_{\mathfrak{M}}} \subseteq \mathrm{SE}(\sigma)$ y **(Kh-2)** $\mathrm{R}_\sigma([\![\psi]\!]^{\mathfrak{L}_{\mathfrak{M}}}) \subseteq [\![\varphi]\!]^{\mathfrak{L}_{\mathfrak{M}}}$, y por ende $w \in [\![\mathsf{Kh}(\psi,\varphi)]\!]^{\mathfrak{L}_{\mathfrak{M}}}$. La dirección $(\supseteq)$ es análoga.

2. Dado que $\mathfrak{L}$ es un LTS por hipótesis y U está bien definido (Observación 4.1), $\mathfrak{M}$ es un LTS^U y por la Definición 6.5, $\mathfrak{M} \in \mathbf{M_{NU}}$. Para demostrar $[\![\varphi]\!]^{\mathfrak{L}} = [\![\varphi]\!]^{\mathfrak{M}_{\mathfrak{L}}}$, los casos para proposiciones atómicas y operadores booleanos se cumplen con el mismo razonamiento que en el item anterior. Para el caso de Kh, $(\subseteq)$ $w \in [\![\mathsf{Kh}(\psi,\varphi)]\!]^{\mathfrak{L}}$ implica que existe un $\sigma \in \mathsf{Act}^*$ tal que **(Kh-1)** $[\![\psi]\!]^{\mathfrak{L}} \subseteq \mathrm{SE}(\sigma)$ y **(Kh-2)** $\mathrm{R}_\sigma([\![\psi]\!]^{\mathfrak{L}}) \subseteq [\![\varphi]\!]^{\mathfrak{L}}$. Por definición de $\mathfrak{M}_{\mathfrak{L}}$, se tiene que $\{\sigma\} \in \mathrm{U}$.

   Por un lado, aplicando hipótesis inductiva, $[\![\psi]\!]^{\mathfrak{M}_{\mathfrak{L}}} = [\![\psi]\!]^{\mathfrak{L}}$. Por **(Kh-1)** y la definición de SE para conjuntos de planes, $[\![\psi]\!]^{\mathfrak{L}} \subseteq \mathrm{SE}(\sigma) = \mathrm{SE}(\{\sigma\})$. Con esto, $[\![\psi]\!]^{\mathfrak{M}_{\mathfrak{L}}} \subseteq \mathrm{SE}(\{\sigma\})$.

   Por otro lado, usando la hipótesis inductiva y la definición de R para conjuntos de planes, $\mathrm{R}_{\{\sigma\}}([\![\psi]\!]^{\mathfrak{M}_{\mathfrak{L}}}) = \mathrm{R}_{\{\sigma\}}([\![\psi]\!]^{\mathfrak{L}}) = \mathrm{R}_\sigma([\![\psi]\!]^{\mathfrak{L}})$. Luego, por **(Kh-2)**, $\mathrm{R}_\sigma([\![\psi]\!]^{\mathfrak{L}}) \subseteq [\![\varphi]\!]^{\mathfrak{L}} = [\![\varphi]\!]^{\mathfrak{M}_{\mathfrak{L}}}$. Con esto, $\mathrm{R}_{\{\sigma\}}([\![\psi]\!]^{\mathfrak{M}_{\mathfrak{L}}}) \subseteq [\![\varphi]\!]^{\mathfrak{M}_{\mathfrak{L}}}$.

   Por ende, **(Kh-1)** $[\![\psi]\!]^{\mathfrak{M}_{\mathfrak{L}}} \subseteq \mathrm{SE}(\{\sigma\})$ y **(Kh-2)** $\mathrm{R}_{\{\sigma\}}([\![\psi]\!]^{\mathfrak{M}_{\mathfrak{L}}}) \subseteq [\![\varphi]\!]^{\mathfrak{M}_{\mathfrak{L}}}$, con esto $w \in [\![\mathsf{Kh}(\psi,\varphi)]\!]^{\mathfrak{M}_{\mathfrak{L}}}$. La otra dirección $(\supseteq)$ es análoga.

□

Esta correspondencia, mostrando que todo LTS tiene un LTS^U equivalente en $\mathbf{M_{NU}}$ y viceversa, nos da el siguiente resultado de correctitud y completitud fuerte.

Teorema 6.3. *El sistema axiomático $\mathcal{L}_{\mathsf{Kh}}^{\mathrm{LTS}}$ (Tabla 3.1) es correcto y fuertemente completo para L_{Kh} con respecto a la clase $\mathbf{M_{NU}}$.*

Demostración. Para correctitud, consideremos los bloques en la Tabla 3.1. Para el primero, el Teorema 5.1 muestra que estos axiomas y reglas son correctas para todo LTS^U, y por lo tanto es es correcto para los modelos de la clase $\mathbf{M_{NU}}$. Para el segundo, el Ítem 1 de la Proposición 6.4 muestra que cada modelo en $\mathbf{M_{NU}}$ es equivalente a un LTS de la lógica L_{Kh}. Usando el Teorema 3.1 se tiene que los axiomas son correctos.

Para demostrar completitud fuerte de $\mathcal{L}_{\mathsf{Kh}}^{\mathrm{LTS}}$ sobre la clase $\mathbf{M_{NU}}$, es necesario mostrar que todo Γ conjunto consistente de fórmulas en $\mathcal{L}_{\mathsf{Kh}}^{\mathrm{LTS}}$ es satisfacible en algún modelo en $\mathbf{M_{NU}}$. Teniendo en cuenta que la Observación 5.2 también se puede aplicar para L_{Kh}, Γ puede ser extendido a un conjunto maximal consistente Γ'. Usando el "truth lemma" de L_{Kh} en [118, Lema 1], existe un LTS $\mathfrak{L}^{\Gamma'}$ tal que $\mathfrak{L}^{\Gamma'}, \Gamma' \models \Gamma$ (notar que los estados en el modelo canónico son conjuntos maximales consistentes). Por el Ítem 2 de la Proposición 6.4, se tiene un LTS^U $\mathfrak{M}_{\mathfrak{L}^{\Gamma'}}$, tal que $\mathfrak{M}_{\mathfrak{L}^{\Gamma'}}, \Gamma' \models \Gamma$. Más aún, $\mathfrak{M}_{\mathfrak{L}^{\Gamma'}}$ está en $\mathbf{M_{NU}}$. □

6.2.2. LTS^Us activos y composicionales

Si bien la clase de modelos $\mathbf{M_{NU}}$ es suficiente para probar que la semántica basada en LTS^Us generaliza a la basada en LTSs, la misma resulta ser demasiado simple. Para ello, en esta sección se define una clase más general de modelos que se comporten como LTSs, basándonos en lo que los axiomas EMP y COMPKh requieren: ***(1)*** un conjunto de planes que se comporten como ϵ y ***(2)*** clausura de U bajo composición de conjuntos de planes. Además, se demuestra que la correspondencia de la sección anterior se cumple para esta clase. Para esto, se define la siguiente noción de composición.

Definición 6.6. Sea $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \mathrm{U}, \mathrm{V} \rangle$ un LTS^U. La *composición* de $\pi_1, \pi_2 \in 2^{\mathsf{Act}^*}$ es el conjunto de planes $\pi_1\pi_2 \in 2^{\mathsf{Act}^*}$ tal que

$$\pi_1\pi_2 := \{\sigma_1\sigma_2 \in \mathsf{Act}^* \mid \sigma_1 \in \pi_1 \text{ y } \sigma_2 \in \pi_2\}.$$

Ejemplo 6.3. Sean, $\pi_1 = \{a, b\}$, $\pi_2 = \{ee\}$ y $\pi_3 = \{f, g\}$, los conjuntos $\pi_1\pi_2$ y $\pi_1\pi_2\pi_3$ son $\{aee, bee\}$ y $\{aeef, beef, aeeg, beeg\}$ respectivamente. □

Con esto, podemos definir formalmente la clase de modelos activos y composicionales.

Definición 6.7. Un LTS^U $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \mathrm{U}, \mathrm{V} \rangle$ es:

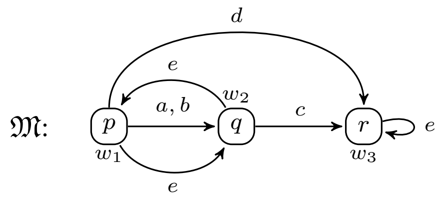
- *activo* si y sólo si existe un $\pi \in \mathrm{U}$ tal que $\mathrm{SE}(\pi) = \mathrm{S}$ y que, para todo $v, u \in \mathrm{S}$, $u \in \mathrm{R}_\pi(v)$ implica $\mathfrak{M}, v \leftrightarrows_{\mathsf{L}_{\mathsf{Kh}_i}} \mathfrak{M}, u$.
- *composicional* si y sólo si para todo $\pi_1, \pi_2 \in \mathrm{U}$ existe $\pi \in \mathrm{U}$ tal que para todo $w \in \mathrm{SE}(\pi_1\pi_2)$, ***(1)*** $w \in \mathrm{SE}(\pi)$, y ***(2)*** $\mathrm{R}_{\pi_1\pi_2}(w) = \mathrm{R}_\pi(w)$.

Con esto se define la clase de modelos LTS^Us activos y composicionales (**AC**):

$$\mathbf{M_{AC}} := \{\mathfrak{M} \mid \mathfrak{M} \text{ es un } \mathrm{LTS}^U \text{ activo y composicional}\}.$$

Por un lado, ser activo garantiza que haya un conjunto de planes que emule lo que el plan vacío ϵ puede realizar en un LTS utilizando la noción de bisimulaciones. Por otro lado, composicionalidad garantiza que U es cerrado bajo una noción conveniente de composición de conjuntos de planes. Dicha conveniencia es que la alcanzabilidad ($\mathrm{R}_{\pi_1\pi_2}(w) = \mathrm{R}_\pi(w)$) sólo se cumple para aquellos estados en los que la composición es fuertemente ejecutable ($w \in \mathrm{SE}(\pi_1\pi_2)$). Además, sólo en estos estados el plan "compositor" es fuertemente ejecutable ($w \in \mathrm{SE}(\pi)$). Notar que la condición de alcanzabilidad es más relajada que la vista en [10], en la cual se requería la igualdad en todos los estados, independientemente de si $\pi_1\pi_2$ era fuertemente ejecutable o no en estos ($\mathrm{R}_{\pi_1\pi_2} = \mathrm{R}_\pi$). A pesar de esta flexibilización, podemos lograr los mismos resultados vistos en [10]. Específicamente, que esta clase de modelos también captura la semántica basada en LTSs. Para ilustrar la definición anterior, consideremos el siguiente ejemplo.

Ejemplo 6.4. Sea $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \mathrm{U}, \mathrm{V}\rangle$ el siguiente LTS^U con $\mathrm{U} = \{\{a, b\}, \{c\}, \{d\}, \{ee\}\}$



Se tiene que $\mathfrak{M}$ es activo dado que el plan $\{ee\}$ cumple con las condiciones. Para verificar que este modelo es composicional, se tiene la siguiente tabla.

π_1	π_1	π	$\mathrm{SE}(\pi_1\pi_2)$	$\mathrm{SE}(\pi)$	$\mathrm{R}_{\pi_1\pi_2}(w)$	$\mathrm{R}_{\pi}(w)$
$\{a, b\}$	$\{c\}$	$\{d\}$	$\{w_1\}$	$\{w_1\}$	$w_1 \to \{w_3\}$	$w_1 \to \{w_3\}$
	$\{ee\}$	$\{a, b\}$	$\{w_1\}$	$\{w_1\}$	$w_1 \to \{w_2\}$	$w_1 \to \{w_2\}$
	$\{a, b\}$	$\{ee\}$	$\emptyset$	S	–	–
	$\{d\}$	$\{ee\}$				
$\{c\}$	$\{ee\}$	$\{c\}$	$\{w_2\}$	$\{w_2\}$	$w_2 \to \{w_3\}$	$w_2 \to \{w_3\}$
	$\{a, b\}$	$\{ee\}$	$\emptyset$	S	–	–
	$\{c\}$	$\{ee\}$				
	$\{d\}$	$\{ee\}$				
$\{d\}$	$\{ee\}$	$\{d\}$	$\{w_1\}$	$\{w_1\}$	$w_1 \to \{w_3\}$	$w_1 \to \{w_3\}$
	$\{a, b\}$	$\{ee\}$	$\emptyset$	S	–	–
	$\{c\}$					
	$\{d\}$					
$\{ee\}$	$\{a, b\}$	$\{a, b\}$	$\{w_1\}$	$\{w_1\}$	$w_1 \to \{w_2\}$	$w_1 \to \{w_2\}$
	$\{c\}$	$\{c\}$	$\{w_2\}$	$\{w_2\}$	$w_2 \to \{w_3\}$	$w_2 \to \{w_3\}$
	$\{d\}$	$\{d\}$	$\{w_1\}$	$\{w_1\}$	$w_1 \to \{w_3\}$	$w_1 \to \{w_3\}$
	$\{ee\}$	$\{ee\}$	S	S	$w_i \to \{w_i\}$	$w_i \to \{w_i\}$

En la misma, se listan exhaustivamente todas las combinaciones de pares de conjuntos de planes $\pi_1, \pi_2 \in \mathrm{U}$, junto con su plan compositor π, y sus correspondientes conjuntos SE. En la sexta y séptima columnas, se tienen los conjuntos $\mathrm{R}_{\pi_1\pi_2}(w)$ y $\mathrm{R}_{\pi}(w)$ para todo $w \in \mathrm{SE}(\pi_1\pi_2)$. Con esto, $\mathfrak{M} \in \mathbf{M_{AC}}$.

Se puede ver además que $\mathsf{Kh}(p, q)$ y $\mathsf{Kh}(q, r)$ se cumplen en este modelo, siendo $\{a, b\}$ y $\{c\}$ sus respectivos testigos. Luego, su plan compositor $\{d\}$ (ver la primera línea de la tabla) es el testigo apropiado para $\mathsf{Kh}(p, r)$. Por otro lado, como $\mathsf{A}(q \to \neg p)$ es una propiedad del modelo, fácilmente $\mathsf{Kh}(q, \neg p)$ también lo es, siendo $\{ee\}$ el plan testigo (que se comporta como el plan vacío ϵ). □

Trivialmente, se tiene la siguiente contención de clases de modelos entre $\mathbf{M_{NU}}$ y $\mathbf{M_{AC}}$.

Proposición 6.5. $\mathbf{M_{NU}} \subseteq \mathbf{M_{AC}}$.

Demostración. Sea $\mathfrak{M} \in \mathbf{M_{NU}}$, $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \mathrm{U}, \mathrm{V}\rangle$ con $\mathrm{U} = \{\{\sigma\} \mid \sigma \in \mathsf{Act}^*\}$. $\mathfrak{M}$ es activo pues existe $\pi = \{\epsilon\} \in \mathrm{U}$ tal que $\mathrm{SE}(\epsilon) = \mathrm{S}$ y $v \in \mathrm{R}_\epsilon(u)$ implica $u = v$ y $\mathfrak{M}, u \leftrightarroweq_{\mathsf{L}_{\mathsf{Kh}_i}} \mathfrak{M}, v$.

Sean $\pi_1, \pi_2 \in \mathrm{U}$ tales que $\pi_i = \{\sigma_i\}$ para $i = 1, 2$, entonces existe $\pi = \pi_1\pi_2 = \{\sigma_1\sigma_2\} \in \mathrm{U}$ tal que $\mathrm{SE}(\pi_1\pi_2) = \mathrm{SE}(\pi)$ y $\mathrm{R}_{\pi_1\pi_2} = \mathrm{R}_\pi$. Por lo tanto, $\mathfrak{M}$ es composicional. □

El siguiente lema establece que los requisitos de la composicionalidad se pueden generalizar a un número arbitrario de conjuntos de planes.

Lema 6.1. *Sea $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \mathrm{U}, \mathrm{V}\rangle$ un LTS^U composicional, y $\pi_1, \ldots, \pi_k \in \mathrm{U}$ (con $k \geq 2$). Entonces existe $\pi \in \mathrm{U}$ tal que para todo $w \in \mathrm{SE}(\pi_1 \cdots \pi_k)$, **(1)** $w \in \mathrm{SE}(\pi)$ y **(2)** $\mathrm{R}_{\pi_1\cdots\pi_k}(w) = \mathrm{R}_\pi(w)$.*

Demostración. La existencia de π se demostrará por inducción en $k \geq 2$. El caso base para $k = 2$ es directo por definición. Sean $\pi_1, \ldots, \pi_k, \pi_{k+1} \in \mathrm{U}$ conjuntos de planes. Por hipótesis inductiva, existe $\pi' \in \mathrm{U}$ tal que para todo $w \in \mathrm{SE}(\pi_1 \cdots \pi_k)$,

- $w \in \mathrm{SE}(\pi')$ y
- $\mathrm{R}_{\pi_1\cdots\pi_k}(w) = \mathrm{R}_{\pi'}(w)$.

Usando la definición de composicionalidad, existe $\pi \in \mathrm{U}$ tal que para todo $w \in \mathrm{SE}(\pi'\pi_{k+1})$,

- $w \in \mathrm{SE}(\pi)$ y
- $\mathrm{R}_{\pi'\pi_{k+1}}(w) = \mathrm{R}_{\pi}(w)$.

Veamos que π cumple con las condiciones. Sea $w \in \mathrm{SE}(\pi_1 \cdots \pi_{k+1})$, para el Ítem ***(1)***, por definición de SE, $w \in \mathrm{SE}(\pi_1 \cdots \pi_k)$ y para todo $v \in \mathrm{R}_{\pi_1\cdots\pi_k}(w)$ se tiene que $v \in \mathrm{SE}(\pi_{k+1})$. Utilizando las propiedades para π', $w \in \mathrm{SE}(\pi')$ y para todo $v \in \mathrm{R}_{\pi'}(w)$ se tiene que $v \in \mathrm{SE}(\pi_{k+1})$. Nuevamente, por definición de SE, $w \in \mathrm{SE}(\pi_1\pi')$. Con esto, $w \in \mathrm{SE}(\pi)$.

Para el Ítem ***(2)***, sea $v \in \mathrm{R}_{\pi_1\cdots\pi_{k+1}}(w)$, existe $u \in \mathrm{R}_{\pi_1\cdots\pi_k}(w)$ tal que $v \in \mathrm{R}_{\pi_{k+1}}(u)$. Por el párrafo anterior, $w \in \mathrm{SE}(\pi_1 \cdots \pi_k) \subseteq \mathrm{SE}(\pi')$ y $\mathrm{R}_{\pi_1\cdots\pi_k}(w) = \mathrm{R}_{\pi'}(w)$. Por lo tanto, $v \in \mathrm{R}_{\pi_1\cdots\pi_{k+1}}(w)$ si y sólo si $v \in \mathrm{R}_{\pi'\pi_{k+1}}(w)$. Como $w \in \mathrm{SE}(\pi_1\pi') \subseteq \mathrm{SE}(\pi)$, $u \in \mathrm{R}_{\pi'\pi_{k+1}}(w)$ si y sólo si $u \in \mathrm{R}_{\pi}(w)$. Con esto, $\mathrm{R}_{\pi_1\cdots\pi_{k+1}}(w) = \mathrm{R}_{\pi}(w)$. $\square$

Con estas propiedades, se puede mostrar que para todo LTS existe un LTS^U equivalente en $\mathbf{M_{AC}}$ y viceversa. Primero, se tiene el mapeo de LTS^Us a LTSs.

Proposición 6.6. *Sea* $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \mathrm{U}, \mathrm{V} \rangle$ *un* LTS^U *en* $\mathbf{M_{AC}}$ *y* $\mathsf{Act}' := \{a_\pi \mid \pi \in \mathrm{U}\}$*, se define el* LTS $\mathfrak{L}_{\mathfrak{M}} = \langle \mathrm{S}, \mathrm{R}, \mathrm{V} \rangle$ *tomando* $\mathrm{R}_{a_\pi} := \{(w, v) \in \mathrm{R}_\pi \mid w \in \mathrm{SE}(\pi)\}$ *(es decir, las acciones básicas en* $\mathfrak{L}_{\mathfrak{M}}$ *corresponden a conjuntos de planes fuertemente ejecutables en* $\mathfrak{M}$*). Entonces* $[\![\varphi]\!]^{\mathfrak{M}} = [\![\varphi]\!]^{\mathfrak{L}_{\mathfrak{M}}}$ *para todo* $\varphi \in \mathsf{L_{Kh}}$.

Demostración. Es directo que $\mathfrak{L}_{\mathfrak{M}}$ es un LTS. El resto de la prueba se hace por inducción estructural en las fórmulas en $\mathsf{L_{Kh}}$. Los casos base y de operadores booleanos son directos, por lo que queda el caso para las fórmulas de la forma $\mathsf{Kh}(\psi, \varphi)$.

Para esto, la siguiente propiedad será útil: para todo $\pi \in \mathrm{U}$, $\mathrm{SE}(\pi) = \mathrm{SE}(a_\pi)$ y para todo $w \in \mathrm{SE}(\pi)$, $\mathrm{R}_\pi(w) = \mathrm{R}_{a_\pi}(w)$. En efecto, **($\subseteq$)** si $w \in \mathrm{SE}(\pi)$ entonces existe $v \in \mathrm{S}$ tal que $(w, v) \in \mathrm{R}_\pi$. Por definición de R_{a_π}, $(w, v) \in \mathrm{R}_{a_\pi}$ y como a_π es una acción básica, $w \in \mathrm{SE}(a_\pi)$. Más aún, **($\supseteq$)** si $w \in \mathrm{SE}(a_\pi)$ entonces hay $v \in \mathrm{S}$ tal que $(w, v) \in \mathrm{R}_{a_\pi}$. Por definición de R_{a_π}, $w \in \mathrm{SE}(\pi)$. La segunda parte de la propiedad es directa de la definición de R_{a_π}.

Esto se puede generalizar a $\mathrm{SE}(\pi_1 \cdots \pi_k) = \mathrm{SE}(a_{\pi_1} \cdots a_{\pi_k})$ y que para todo $w \in \mathrm{SE}(\pi_1 \cdots \pi_k)$, $\mathrm{R}_{\pi_1\cdots\pi_k}(w) = \mathrm{R}_{a_{\pi_1}\cdots a_{\pi_k}}(w)$ con $k \geq 1$. El caso base se demostró en el párrafo anterior. Sea $w \in \mathrm{SE}(\pi_1 \cdots \pi_{k+1})$, usando la definición de ejecutabilidad fuerte, $w \in \mathrm{SE}(\pi_1 \cdots \pi_k)$ y para todo $v \in \mathrm{R}_{\pi_1\cdots\pi_k}(w)$, se tiene que $v \in \mathrm{SE}(\pi_{k+1})$. Por hipótesis inductiva y el caso base, $w \in \mathrm{SE}(a_{\pi_1} \cdots a_{\pi_k})$ y para todo $v \in \mathrm{R}_{a_{\pi_1}\cdots a_{\pi_k}}(w)$, se tiene que $v \in \mathrm{SE}(a_{\pi_{k+1}})$. Que es equivalente a $w \in \mathrm{SE}(a_{\pi_1} \cdots a_{\pi_{k+1}})$. Para la segunda parte, $(w, z) \in \mathrm{R}_{\pi_1\cdots\pi_{k+1}}$ se cumple si y sólo si existe $v \in \mathrm{S}$ tal que $(w, v) \in \mathrm{R}_{\pi_1\cdots\pi_k}$ y $(v, z) \in \mathrm{R}_{\pi_{k+1}}$. Como $w \in \mathrm{SE}(\pi_1 \cdots \pi_{k+1})$, entonces $w \in \mathrm{SE}(\pi_1 \cdots \pi_k)$ y para todo $v \in \mathrm{R}_{\pi_1\cdots\pi_k}(w)$, se tiene que $v \in \mathrm{SE}(\pi_{k+1})$. Por lo tanto, existe $v \in \mathrm{S}$ tal que $(w, v) \in \mathrm{R}_{a_{\pi_1}\cdots a_{\pi_k}}$ y $(v, z) \in \mathrm{R}_{a_{\pi_{k+1}}}$. Con esto, $(w, z) \in \mathrm{R}_{a_{\pi_1}\cdots a_{\pi_{k+1}}}$. La vuelta es análoga. Procedamos a demostrar el caso Kh.

($\subseteq$) Sea $w \in [\![\mathsf{Kh}(\psi, \varphi)]\!]^{\mathfrak{M}}$, entonces existe $\pi \in \mathrm{U}$ tal que cumple

(Kh-1) $[\![\psi]\!]^{\mathfrak{M}} \subseteq \mathrm{SE}(\pi)$ y **(Kh-2)** $\mathrm{R}_\pi([\![\psi]\!]^{\mathfrak{M}}) \subseteq [\![\varphi]\!]^{\mathfrak{M}}$.

Se demostrará que $w \in [\![\mathsf{Kh}(\psi, \varphi)]\!]^{\mathfrak{L}_{\mathfrak{M}}}$ usando $a_\pi \in \mathsf{Act}'$ como testigo. Primero, sea $v \in [\![\psi]\!]^{\mathfrak{L}_{\mathfrak{M}}}$, por hipótesis inductiva, $v \in [\![\psi]\!]^{\mathfrak{M}}$, y, por **(Kh-1)**, $v \in \mathrm{SE}(\pi)$. Por la propiedad anterior, $v \in \mathrm{SE}(a_\pi)$. Con esto, $[\![\psi]\!]^{\mathfrak{L}_{\mathfrak{M}}} \subseteq \mathrm{SE}(a_\pi)$. Segundo, sea $u \in \mathrm{R}_{a_\pi}([\![\psi]\!]^{\mathfrak{L}_{\mathfrak{M}}})$, por hipótesis inductiva, $u \in \mathrm{R}_{a_\pi}([\![\psi]\!]^{\mathfrak{M}})$ y, por definición de R, $u \in \mathrm{R}_\pi([\![\psi]\!]^{\mathfrak{M}})$. Con esto, por **(Kh-2)**, $u \in [\![\varphi]\!]^{\mathfrak{M}}$ y por hipótesis inductiva $u \in [\![\varphi]\!]^{\mathfrak{L}_{\mathfrak{M}}}$. Por lo tanto, $\mathrm{R}_{a_\pi}([\![\psi]\!]^{\mathfrak{L}_{\mathfrak{M}}}) \subseteq [\![\varphi]\!]^{\mathfrak{L}_{\mathfrak{M}}}$. De estas dos partes, se tiene que $w \in [\![\mathsf{Kh}(\psi, \varphi)]\!]^{\mathfrak{L}_{\mathfrak{M}}}$.

($\supseteq$) Sea $w \in [\![\mathsf{Kh}(\psi, \varphi)]\!]^{\mathfrak{L}_{\mathfrak{M}}}$, entonces existe $\sigma \in (\mathsf{Act}')^*$ tal que cumple

(Kh-1) $[\![\psi]\!]^{\mathfrak{L}_{\mathfrak{M}}} \subseteq \mathrm{SE}(\sigma)$ y **(Kh-2)** $\mathrm{R}_\sigma([\![\psi]\!]^{\mathfrak{L}_{\mathfrak{M}}}) \subseteq [\![\varphi]\!]^{\mathfrak{L}_{\mathfrak{M}}}$.

Hay dos casos principales. Si $\sigma = \epsilon$, $[\![\psi]\!]^{\mathfrak{L}_{\mathfrak{M}}} \subseteq [\![\varphi]\!]^{\mathfrak{L}_{\mathfrak{M}}}$, que por hipótesis inductiva equivale a $[\![\psi]\!]^{\mathfrak{M}} \subseteq [\![\varphi]\!]^{\mathfrak{M}}$. Como $\mathfrak{M}$ es activo, hay un $\pi \in \mathrm{U}$ tal que $\mathrm{SE}(\pi) = \mathrm{S}$ y para todo $v, u \in \mathrm{S}$,

$u \in \mathrm{R}_\pi(v)$ implica $\mathfrak{M}, v \leftrightarrows_{\mathsf{L_{Kh_i}}} \mathfrak{M}, w$. Sea $u \in \mathrm{R}_\pi([\![\psi]\!]^{\mathfrak{M}})$, por definición existe $v \in [\![\psi]\!]^{\mathfrak{M}} \subseteq [\![\varphi]\!]^{\mathfrak{M}}$ tal que $u \in \mathrm{R}_\pi(v)$. Luego, $\mathfrak{M}, v \leftrightarrows_{\mathsf{L_{Kh_i}}} \mathfrak{M}, u$ y usando el Teorema 6.1, $\mathfrak{M}, v \leftrightsquigarrow_{\mathsf{L_{Kh_i}}} \mathfrak{M}, u$. Por lo tanto, $u \in [\![\varphi]\!]^{\mathfrak{M}}$ y π es el testigo apropiado.

Si en cambio, $\sigma \neq \epsilon$, se tiene que $\sigma = a_{\pi_1} \cdots a_{\pi_k}$ con $a_{\pi_i} \in \mathsf{Act}'$ (por ende, $\pi_i \in \mathrm{U}$). Hay dos situaciones. Si $[\![\psi]\!]^{\mathfrak{L}_{\mathfrak{M}}} = \emptyset$ entonces, por hipótesis inductiva, $[\![\psi]\!]^{\mathfrak{M}} = \emptyset$. Luego cualquier conjunto de planes $\pi \in \mathrm{U}$ funciona como testigo. Caso contrario, $[\![\psi]\!]^{\mathfrak{L}_{\mathfrak{M}}} \neq \emptyset$ y, como $\mathfrak{M}$ es composicional, por el Lema 6.1 existe $\pi \in \mathrm{U}$ tal que para todo $v \in \mathrm{SE}(\pi_1 \cdots \pi_k)$, $v \in \mathrm{SE}(\pi)$ y $\mathrm{R}_{\pi_1 \cdots \pi_k}(v) = \mathrm{R}_\pi(v)$.

Para la primera condición de Kh, sea $v \in [\![\psi]\!]^{\mathfrak{M}}$. Por hipótesis inductiva, $v \in [\![\psi]\!]^{\mathfrak{L}_{\mathfrak{M}}}$, y por **(Kh-1)**, $v \in \mathrm{SE}(a_{\pi_1} \cdots a_{\pi_k})$. Usando las propiedades demostradas al principio, $v \in \mathrm{SE}(\pi_1 \cdots \pi_k)$. Por lo tanto, $v \in \mathrm{SE}(\pi)$. Para la segunda condición de Kh, sean $v, u \in \mathrm{S}$ tales que $v \in [\![\psi]\!]^{\mathfrak{M}}$ y $(v, u) \in \mathrm{R}_\pi$. Como $v \in \mathrm{SE}(\pi_1 \cdots \pi_k)$, entonces $(v, u) \in \mathrm{R}_{\pi_1 \cdots \pi_k}$. Considerando además que $\mathrm{SE}(\pi_1 \cdots \pi_k) = \mathrm{SE}(a_{\pi_1} \cdots a_{\pi_k})$, se tiene que $(v, u) \in \mathrm{R}_{a_{\pi_1} \cdots a_{\pi_k}}$. Por hipótesis inductiva, $v \in [\![\psi]\!]^{\mathfrak{L}_{\mathfrak{M}}}$ y por **(Kh-2)**, $u \in [\![\varphi]\!]^{\mathfrak{L}_{\mathfrak{M}}}$. Nuevamente, aplicando hipótesis inductiva, $u \in [\![\varphi]\!]^{\mathfrak{M}}$. Por lo tanto, $\mathrm{R}_\pi([\![\psi]\!]^{\mathfrak{M}}) \subseteq [\![\varphi]\!]^{\mathfrak{M}}$. De las dos partes, se tiene que $w \in [\![\mathsf{Kh}(\psi, \varphi)]\!]^{\mathfrak{M}}$. □

Antes de probar la otra dirección, es necesario dar la definición de una relación sobre todos los planes considerados. Dicha relación permitira agrupar aquellos planes que se comporten de una manera similar.

Definición 6.8. Sea $\mathfrak{L} = \langle \mathrm{S}, \mathrm{R}, \mathrm{V} \rangle$ un LTS, se define la siguiente relación $\bowtie$ sobre Act^*.

$$\sigma_1 \bowtie \sigma_2 \text{ sii } \mathrm{SE}(\sigma_1) = \mathrm{SE}(\sigma_2) \text{ y para todo } w \in \mathrm{SE}(\sigma_1), \text{ se tiene que } \mathrm{R}_{\sigma_1}(w) = \mathrm{R}_{\sigma_2}(w).$$

En otras palabras, $\bowtie$ separa los planes en grupos que actúan similarmente en términos de ejecutabilidad fuerte y alcanzabilidad. De esta relación se pueden deducir las siguientes propiedades.

Proposición 6.7. *Sean* $\mathfrak{L} = \langle \mathrm{S}, \mathrm{R}, \mathrm{V} \rangle$ *un* LTS, $\sigma \in \mathsf{Act}^*$ *y* $\pi_\sigma = \{\sigma' \in \mathsf{Act}^* \mid \sigma \bowtie \sigma'\}$. *Entonces*

- $\bowtie$ *es una relación de equivalencia;*
- $\mathrm{SE}(\sigma) = \mathrm{SE}(\pi_\sigma)$ *y para todo* $w \in \mathrm{SE}(\sigma)$, *se tiene que* $\mathrm{R}_\sigma(w) = \mathrm{R}_{\pi_\sigma}(w)$*;*
- *para todo* $\sigma_1, \sigma'_1, \sigma_2, \sigma'_2 \in \mathsf{Act}^*$, *si* $\sigma_1 \bowtie \sigma'_1$ *y* $\sigma_2 \bowtie \sigma'_2$, *entonces* $\sigma_1\sigma_2 \bowtie \sigma'_1\sigma'_2$*;*
- *para todo* $\sigma_1, \sigma_2 \in \mathsf{Act}^*$, $\pi_{\sigma_1}\pi_{\sigma_2} \subseteq \pi_{\sigma_1\sigma_2}$ *y por lo tanto* $\mathrm{SE}(\pi_{\sigma_1\sigma_2}) \subseteq \mathrm{SE}(\pi_{\sigma_1}\pi_{\sigma_2})$*; y*
- *para todo* $k \geq 2$ *y* $\sigma_1, \ldots, \sigma_k \in \mathsf{Act}^*$, $\mathrm{SE}(\pi_{\sigma_1} \cdots \pi_{\sigma_k}) = \mathrm{SE}(\pi_{\sigma_1 \cdots \sigma_k})$ *y para todo* $w \in \mathrm{SE}(\pi_{\sigma_1} \cdots \pi_{\sigma_k})$, *se tiene que* $\mathrm{R}_{\pi_{\sigma_1} \cdots \pi_{\sigma_k}}(w) = \mathrm{R}_{\pi_{\sigma_1 \cdots \sigma_k}}(w)$.

Demostración. La demostración de los primeros dos ítems es directa usando las definiciones de $\bowtie$ y π_σ. Procedamos con los últimos tres.

- Sea $w \in \mathrm{SE}(\sigma_1\sigma_2)$, por definición de SE, esto equivale a que $w \in \mathrm{SE}(\sigma_1)$ y para todo $v \in \mathrm{R}_{\sigma_1}(w)$, se tiene que $v \in \mathrm{SE}(\sigma_2)$. Por definición de $\bowtie$, $w \in \mathrm{SE}(\sigma'_1)$ y para todo $v \in \mathrm{R}_{\sigma'_1}(w)$, se tiene que $v \in \mathrm{SE}(\sigma'_2)$. Nuevamente, por definición de SE, $w \in \mathrm{SE}(\sigma'_1\sigma'_2)$. Sean $w \in \mathrm{SE}(\sigma_1\sigma_2)$ y $v \in \mathrm{R}_{\sigma_1\sigma_2}(w)$, entonces existe $u \in \mathrm{R}_{\sigma_1}(w)$ tal que $v \in \mathrm{R}_{\sigma_2}(u)$. Como $w \in \mathrm{SE}(\sigma_1) = \mathrm{SE}(\sigma'_1)$ y $v \in \mathrm{SE}(\sigma_2) = \mathrm{SE}(\sigma'_2)$, entonces $u \in \mathrm{R}_{\sigma'_1}(w)$, $v \in \mathrm{R}_{\sigma'_2}(u)$ y con esto $v \in \mathrm{R}_{\sigma'_1\sigma'_2}(w)$. La vuelta es análoga. Con esto, $\sigma_1\sigma_2 \bowtie \sigma'_1\sigma'_2$.
- Sean $\sigma'_1 \in \pi_{\sigma_1}$ y $\sigma'_2 \in \pi_{\sigma_2}$, por definición de estos conjuntos, $\sigma_1 \bowtie \sigma'_1$ y $\sigma_2 \bowtie \sigma'_2$. Por el ítem anterior, $\sigma_1\sigma_2 \bowtie \sigma'_1\sigma'_2$ y con esto $\sigma'_1\sigma'_2 \in \pi_{\sigma_1\sigma_2}$.
- Para el caso $k = 2$, sea $w \in \mathrm{SE}(\pi_{\sigma_1}\pi_{\sigma_2})$, por definición de SE, $w \in \mathrm{SE}(\pi_{\sigma_1})$ y para todo $v \in \mathrm{R}_{\pi_{\sigma_1}}(w)$, se tiene que $v \in \mathrm{SE}(\pi_{\sigma_2})$. Luego, por la definición de π_{σ_1} y π_{σ_2}, $w \in \mathrm{SE}(\sigma_1)$ y para todo $v \in \mathrm{R}_{\sigma_1}(w)$, se tiene que $v \in \mathrm{SE}(\sigma_2)$. Por definición de SE, $w \in \mathrm{SE}(\sigma_1\sigma_2) = \mathrm{SE}(\pi_{\sigma_1\sigma_2})$. Con el ítem anterior se consigue la otra contención. Para la otra propiedad, sea $w \in \mathrm{SE}(\pi_{\sigma_1}\pi_{\sigma_2})$, $v \in \mathrm{R}_{\pi_{\sigma_1}\pi_{\sigma_2}}(w)$ si y sólo si existe $u \in \mathrm{R}_{\pi_{\sigma_1}}(w)$ tal que $v \in \mathrm{R}_{\pi_{\sigma_2}}(w)$. Por definición de π_{σ_1} y π_{σ_2}, esto equivale a que existe $u \in \mathrm{R}_{\sigma_1}(w)$ tal que $v \in \mathrm{R}_{\sigma_2}(w)$. Por lo tanto, $v \in \mathrm{R}_{\sigma_1\sigma_2}(w) = \mathrm{R}_{\pi_{\sigma_1\sigma_2}}(w)$.

Sean $k \geq 2$, $\sigma_1, \dots, \sigma_k, \sigma_{k+1} \in \mathsf{Act}^*$ y $w \in \mathrm{SE}(\pi_{\sigma_1} \cdots \pi_{\sigma_{k+1}})$, por definición de SE, $w \in \mathrm{SE}(\pi_{\sigma_1} \cdots \pi_{\sigma_k})$ y para todo $v \in \mathrm{R}_{\pi_{\sigma_1} \cdots \pi_{\sigma_k}}(w)$, se tiene que $v \in \mathrm{SE}(\pi_{\sigma_{k+1}})$. Por hipótesis inductiva, $w \in \mathrm{SE}(\pi_{\sigma_1 \cdots \sigma_k})$ y para todo $v \in \mathrm{R}_{\pi_{\sigma_1 \cdots \sigma_k}}(w)$, se tiene que $v \in \mathrm{SE}(\pi_{\sigma_{k+1}})$. Nuevamente, por definición de SE, esto equivale a que $w \in \mathrm{SE}(\pi_{\sigma_1 \cdots \sigma_k} \pi_{\sigma_{k+1}})$. Aplicando el caso base, lo anterior es equivalente a que $w \in \mathrm{SE}(\pi_{\sigma_1 \cdots \sigma_{k+1}})$. Para la otra propiedad, sean $w \in \mathrm{SE}(\pi_{\sigma_1} \cdots \pi_{\sigma_{k+1}})$ y $v \in \mathrm{R}_{\pi_{\sigma_1} \cdots \pi_{\sigma_{k+1}}}(w)$, se tiene que existe $u \in \mathrm{R}_{\pi_{\sigma_1} \cdots \pi_{\sigma_k}}(w)$ tal que $v \in \mathrm{R}_{\pi_{\sigma_{k+1}}}(u)$. Por definición de SE e hipótesis inductiva, equivalentemente existe $u \in \mathrm{R}_{\pi_{\sigma_1 \cdots \sigma_k}}(w)$ tal que $v \in \mathrm{R}_{\pi_{\sigma_{k+1}}}(u)$. Por lo tanto, $v \in \mathrm{R}_{\pi_{\sigma_1 \cdots \sigma_k} \pi_{\sigma_{k+1}}}(w)$. Como $\mathrm{SE}(\pi_{\sigma_1} \cdots \pi_{\sigma_{k+1}}) = \mathrm{SE}(\pi_{\sigma_1 \cdots \sigma_k} \pi_{\sigma_{k+1}})$, aplicando el caso base, $v \in \mathrm{R}_{\pi_{\sigma_1 \cdots \sigma_{k+1}}}(w)$.

$\square$

Con esto, podemos demostrar que, desde un LTS se puede conseguir un LTS^U activo y composicional equivalente.

Proposición 6.8. *Sea* $\mathfrak{L} = \langle \mathrm{S}, \mathrm{R}, \mathrm{V} \rangle$ *un* LTS, *existe un* LTS^U $\mathfrak{M}_{\mathfrak{L}} = \langle \mathrm{S}, \mathrm{R}, \mathrm{U}, \mathrm{V} \rangle$ *tal que* $\mathrm{U} = \{\pi_\sigma \mid \sigma \in \mathsf{Act}^*\}$ *con* $\pi_\sigma = \{\sigma' \in \mathsf{Act}^* \mid \sigma \bowtie \sigma'\}$. *Entonces,*

1. $\mathfrak{M}_{\mathfrak{L}}$ *es activo y composicional* ($\mathfrak{M}_{\mathfrak{L}} \in \mathbf{M_{AC}}$)*;*
2. *para todo* $\varphi \in \mathsf{L_{Kh}}$, $[\![\varphi]\!]^{\mathfrak{L}} = [\![\varphi]\!]^{\mathfrak{M}_{\mathfrak{L}}}$.

Demostración. Para el Ítem 1, notar que $\mathfrak{M}_{\mathfrak{L}}$ es un LTS^U. Como $\mathsf{Act}^* \neq \emptyset$, se tiene que $\mathrm{P}' = \bigcup_{\pi \in \mathrm{U}} \pi$ es no vacío. Sean $\pi_\sigma, \pi_{\sigma'} \in \mathrm{U}$ tales que $\pi_\sigma \neq \pi_{\sigma'}$, como $\bowtie$ es una relación de equivalencia, entonces π_σ y $\pi_{\sigma'}$ son clases de equivalencia. Por lo tanto, $\pi_\sigma \cap \pi_{\sigma'} = \emptyset$. Trivialmente, U no contiene el conjunto vacío.

Para demostrar que es activo, tomamos el conjunto π_ϵ. Por un lado, $\mathrm{SE}(\pi_\epsilon) = \mathrm{SE}(\epsilon) = \mathrm{S}$. Por el otro, dados $v, u \in \mathrm{S}$, tales que $u \in \mathrm{R}_{\pi_\epsilon}(v)$, como $v \in \mathrm{SE}(\epsilon)$, se tiene que $\mathrm{R}_{\pi_\epsilon}(v) = \mathrm{R}_\epsilon(v) = \{v\}$. Con esto, $u = v$ y es directo que $\mathfrak{M}_{\mathfrak{L}}, v \leftrightarroweq_{\mathsf{L_{Kh_i}}} \mathfrak{M}_{\mathfrak{L}}, u$ ($\leftrightarroweq_{\mathsf{L_{Kh_i}}}$ es reflexiva). Luego, $\mathfrak{M}_{\mathfrak{L}}$ es activo.

Para ver que es composicional, sean $\pi_\sigma, \pi_{\sigma'} \in \mathrm{U}$. Se propone el conjunto $\pi_{\sigma\sigma'} \in \mathrm{U}$. Por la Proposición 6.7, $\mathrm{SE}(\pi_\sigma \pi_{\sigma'}) = \mathrm{SE}(\pi_{\sigma\sigma'})$ y para todo $w \in \mathrm{SE}(\pi_\sigma \pi_{\sigma'})$, se tiene que $\mathrm{R}_{\pi_\sigma \pi_{\sigma'}}(w) = \mathrm{R}_{\pi_{\sigma\sigma'}}(w)$. Para el Ítem 2, la prueba es por inducción en las fórmulas. De vuelta, sólo se muestra el caso para $\mathsf{Kh}(\psi, \varphi)$.

($\subseteq$) Sea $w \in [\![\mathsf{Kh}(\psi, \varphi)]\!]^{\mathfrak{L}}$, entonces existe $\sigma \in \mathsf{Act}^*$ tal que

(Kh-1) $[\![\psi]\!]^{\mathfrak{L}} \subseteq \mathrm{SE}(\sigma)$ y **(Kh-2)** $\mathrm{R}_\sigma([\![\psi]\!]^{\mathfrak{L}}) \subseteq [\![\varphi]\!]^{\mathfrak{L}}$.

Si $[\![\psi]\!]^{\mathfrak{M}_{\mathfrak{L}}} = \emptyset$, entonces cualquier $\pi \in \mathrm{U}$ es testigo. Si $[\![\psi]\!]^{\mathfrak{M}_{\mathfrak{L}}} \neq \emptyset$, $\pi_\sigma \in \mathrm{U}$ será el testigo. Para la primera condición de Kh, si $u \in [\![\psi]\!]^{\mathfrak{M}_{\mathfrak{L}}}$ entonces por hipótesis inductiva $u \in [\![\psi]\!]^{\mathfrak{L}}$, y por **(Kh-1)**, $u \in \mathrm{SE}(\sigma)$. Por la Proposición 6.7, $u \in \mathrm{SE}(\pi_\sigma)$. Para la segunda condición de Kh, sea $v \in \mathrm{R}_{\pi_\sigma}([\![\psi]\!]^{\mathfrak{M}_{\mathfrak{L}}})$, se tiene que $v \in \mathrm{R}_{\pi_\sigma}(u)$ para algún $u \in [\![\psi]\!]^{\mathfrak{M}_{\mathfrak{L}}} \subseteq \mathrm{SE}(\pi_\sigma)$. Por la Proposición 6.7, $v \in \mathrm{R}_\sigma(u)$ y por hipótesis inductiva $u \in [\![\psi]\!]^{\mathfrak{L}}$. Por **(Kh-2)** $v \in [\![\varphi]\!]^{\mathfrak{L}}$ y nuevamente por hipótesis inductiva $v \in [\![\varphi]\!]^{\mathfrak{M}_{\mathfrak{L}}}$. Con esto, $\mathrm{R}_\pi([\![\psi]\!]^{\mathfrak{M}_{\mathfrak{L}}}) \subseteq [\![\varphi]\!]^{\mathfrak{M}_{\mathfrak{L}}}$. De las dos condiciones, se tiene que $w \in [\![\mathsf{Kh}(\psi, \varphi)]\!]^{\mathfrak{M}_{\mathfrak{L}}}$.

($\supseteq$) Sea $w \in [\![\mathsf{Kh}(\psi, \varphi)]\!]^{\mathfrak{M}_{\mathfrak{L}}}$, entonces existe $\pi_\sigma \in \mathrm{U}$ tal que

(Kh-1) $[\![\psi]\!]^{\mathfrak{M}_{\mathfrak{L}}} \subseteq \mathrm{SE}(\pi_\sigma)$ y **(Kh-2)** $\mathrm{R}_{\pi_\sigma}([\![\psi]\!]^{\mathfrak{M}_{\mathfrak{L}}}) \subseteq [\![\varphi]\!]^{\mathfrak{M}_{\mathfrak{L}}}$.

Se verá que σ es el testigo. Para la primera condición de Kh, sea $u \in [\![\psi]\!]^{\mathfrak{L}}$, por hipótesis inductiva $u \in [\![\psi]\!]^{\mathfrak{M}_{\mathfrak{L}}}$ y por **(Kh-1)**, $u \in \mathrm{SE}(\pi_\sigma)$. Por la Proposición 6.7, $u \in \mathrm{SE}(\sigma)$. Por lo tanto, $[\![\psi]\!]^{\mathfrak{L}} \subseteq \mathrm{SE}(\sigma)$. Para la segunda condición de Kh, sea $u \in \mathrm{R}_\sigma([\![\psi]\!]^{\mathfrak{L}})$, por hipótesis inductiva $u \in \mathrm{R}_\sigma([\![\psi]\!]^{\mathfrak{M}_{\mathfrak{L}}})$. Luego $u \in \mathrm{R}_\sigma(v)$ para algún $v \in [\![\psi]\!]^{\mathfrak{M}_{\mathfrak{L}}}$. Por **(Kh-1)**, $v \in \mathrm{SE}(\pi_\sigma)$, y por la Proposición 6.7, $u \in \mathrm{R}_{\pi_\sigma}(v)$. Luego, $u \in \mathrm{R}_{\pi_\sigma}([\![\psi]\!]^{\mathfrak{M}_{\mathfrak{L}}})$, y por **(Kh-2)** $u \in [\![\varphi]\!]^{\mathfrak{M}_{\mathfrak{L}}}$. Por hipótesis inductiva, $u \in [\![\varphi]\!]^{\mathfrak{L}}$. Con esto, $\mathrm{R}_\sigma([\![\psi]\!]^{\mathfrak{L}}) \subseteq [\![\varphi]\!]^{\mathfrak{L}}$. Juntando ambas partes, se tiene que $w \in [\![\mathsf{Kh}(\psi, \varphi)]\!]^{\mathfrak{L}}$. $\square$

De estos resultados, podemos demostrar que el sistema axiomático para $\mathsf{L_{Kh}}$ sobre LTS (Tabla 3.1) es correcto y fuertemente completo para $\mathsf{L_{Kh}}$ sobre LTS^Us *activos* y *composicionales.*

Teorema 6.4. *El sistema axiomático* $\mathcal{L}_{\mathsf{Kh}}^{\mathrm{LTS}}$ *(Tabla 3.1) es correcto y fuertemente completo con respecto a la clase* $\mathbf{M_{AC}}$ *de* LTS^U*s activos y composicionales.*

Demostración. Para correctitud, consideremos los bloques en la Tabla 3.1. Para el primero, el Teorema 5.1 muestra que estos axiomas y reglas son correctas para todo LTS^U, y por lo tanto es es correcto para los modelos de la clase $\mathbf{M_{AC}}$. Para el segundo, la Proposición 6.6 muestra que cada modelo en $\mathbf{M_{AC}}$ es equivalente a un LTS de la lógica $\mathsf{L_{Kh}}$. Usando el Teorema 3.1 se tiene que los axiomas son correctos.

Para demostrar completitud fuerte de $\mathcal{L}_{\mathsf{Kh}}^{\text{LTS}}$ sobre la clase $\mathbf{M_{AC}}$, es necesario mostrar que todo Γ conjunto consistente de fórmulas en $\mathcal{L}_{\mathsf{Kh}}^{\text{LTS}}$ es satisfacible en algún modelo en $\mathbf{M_{AC}}$. Teniendo en cuenta que la Observación 5.2 también se puede aplicar para $\mathsf{L_{Kh}}$, Γ puede ser extendido a un conjunto maximal consistente Γ'. Usando el "truth lemma" de $\mathsf{L_{Kh}}$ en [118, Lema 1], existe un LTS $\mathfrak{L}^{\Gamma'}$ tal que $\mathfrak{L}^{\Gamma'}, \Gamma' \models \Gamma$ (notar que los estados en el modelo canónico son conjuntos maximales consistentes). Por la Proposición 6.8, se tiene un LTS^U $\mathfrak{M}_{\mathfrak{L}^{\Gamma'}}$, tal que $\mathfrak{M}_{\mathfrak{L}^{\Gamma'}}, \Gamma' \models \Gamma$. Más aún, $\mathfrak{M}_{\mathfrak{L}^{\Gamma'}}$ está en $\mathbf{M_{AC}}$. $\square$

Resumiendo, mientras que en la Subsección 6.2.1 con $\mathbf{M_{NU}}$ tenemos un agente que tiene todos los planes disponibles y es capaz de distinguirlos entre sí, $\mathbf{M_{AC}}$ es más general y se pide únicamente que los conjuntos de planes considerados "emulen" el comportamiento de los planes en los LTSs. Dado que $\mathcal{L}_{\mathsf{Kh}}^{\text{LTS}}$ es correcta y fuertemente completa con respecto a ambas clases, se tiene que la semántica basada en LTSs puede ser generalizada considerando cualquiera de estas, siendo $\mathbf{M_{AC}}$ la más abarcativa.

Capítulo 7

Propiedades computacionales

> Meditad sobre el porqué, pero más aun sobre el cómo.
>
> Goethe

En este capítulo nos centraremos en el estudio de la complejidad computacional de la lógica $\mathsf{L}_{\mathsf{Kh}_i}$ sobre LTS^Us. Para ello, se utilizarán dos herramientas estándar de la lógica modal trabajadas en [10].

La primera herramienta es definir una noción de filtración con la cual, dada una fórmula y un modelo arbitrario, se obtenga uno finito que preserve la satisfacibilidad de dicha fórmula. Este método de filtración [70, 98, 99] consiste en, dado un modelo arbitrario (potencialmente infinito) $\mathfrak{M}$ que satisface una determinada fórmula φ, colapsar las componentes de $\mathfrak{M}$ con respecto a φ y sus subfórmulas, de manera tal que se elimina la información "repetida". Por ejemplo, dos estados que satisfacen las mismas subfórmulas de φ, pertenecerán al mismo representante (misma clase de equivalencia con respecto a φ). Notar que para la lógica $\mathsf{L}_{\mathsf{Kh}_i}$, debemos ser cuidadosos en seleccionar las relaciones de incertidumbre entre planes. Esto demuestra que el problema de satisfacibilidad para $\mathsf{L}_{\mathsf{Kh}_i}$ es al menos decidible.

La segunda herramienta es definir una función de selección la cual, desde el modelo canónico, extraiga un modelo de tamaño polinomial con respecto al tamaño de la fórmula. Este método de selección [24] consiste en elegir las partes importantes de un modelo para la evaluación de cierta fórmula, y descartar toda la información superflua del mismo. Además, demostraremos que el problema de model checking para $\mathsf{L}_{\mathsf{Kh}_i}$ está en P, es decir, existe un algoritmo que lo resuelve en tiempo polinomial en función del tamaño del modelo y de la fórmula. Con esto se muestra que el problema de satisfacibilidad para $\mathsf{L}_{\mathsf{Kh}_i}$ es NP-completo.

Notar que el segundo método incorpora el resultado de complejidad a diferencia del primero. Sin embargo, el método de selección en nuestro caso funciona específicamente sobre el modelo canónico, mientras que la filtración puede aplicarse sobre un modelo arbitrario. Por lo tanto, la filtración nos permite obtener en la práctica un modelo finito para cualquier fórmula a partir de un modelo arbitrario de la misma.

7.1. Propiedad de modelo finito vía filtraciones

Comenzaremos presentando el método de filtración para obtener modelos finitos. Con este fin, necesitamos colapsar por un lado estados y por el otro planes que capturen la misma información acerca de una fórmula dada. Introducimos entonces dos relaciones que serán importantes para definir una noción de filtración, dado un conjunto de fórmulas Σ y un modelo $\mathfrak{M}$.

Definición 7.1 (Σ-equivalencia). Sean $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \{\mathrm{U}(i)\}_{i \in \mathsf{Agt}}, \mathrm{V} \rangle$ un LTS^U y Σ un conjunto de fórmulas de $\mathsf{L}_{\mathsf{Kh}_i}$ cerrado bajo subfórmulas. Se definen las relaciones $\leftrightsquigarrow_\Sigma \subseteq \mathrm{S} \times \mathrm{S}$ y $\leftrightarrows_\Sigma \subseteq \mathrm{U}_{\mathsf{Agt}} \times \mathrm{U}_{\mathsf{Agt}}$ (con $\mathrm{U}_{\mathsf{Agt}} := \bigcup_{i \in \mathsf{Agt}} \mathrm{U}(i)$) de la siguiente forma:

$$
\begin{array}{ll}
w \leftrightsquigarrow_\Sigma v & \text{sii para todo } \psi \in \Sigma,\ \mathfrak{M}, w \models \psi \text{ sii } \mathfrak{M}, v \models \psi, \\
\pi \leftrightarrows_\Sigma \pi' & \text{sii para todo } i \in \mathsf{Agt} \text{ y } \mathsf{Kh}_i(\psi, \varphi) \in \Sigma,\ \pi \text{ es un testigo para} \\
 & \quad \mathsf{Kh}_i(\psi, \varphi) \text{ en } \mathfrak{M} \text{ sii } \pi' \text{ es un testigo para } \mathsf{Kh}_i(\psi, \varphi) \text{ en } \mathfrak{M}.
\end{array}
$$

Notar que $\rightsquigarrow_\Sigma$, una generalización de $\rightsquigarrow$ en la Definición 6.3 para un conjunto de fórmulas Σ, y $\leftrightarrows_\Sigma$ son una relación de equivalencia sobre S y $\mathrm{U}_{\mathsf{Agt}}$ respectivamente. Para $w \in \mathrm{S}$ (respectivamente $\pi \in \mathrm{U}_{\mathsf{Agt}}$), se usa $[w]_\Sigma$ (respectivamente $[\pi]_\Sigma$) para denotar la clase de equivalencia w (respectivamente π) sobre Σ; es decir,

$$[w]_\Sigma := \{v \in \mathrm{S} \mid w \rightsquigarrow_\Sigma v\}; \qquad [\pi]_\Sigma := \{\pi' \in \mathrm{U}_{\mathsf{Agt}} \mid \pi \leftrightarrows_\Sigma \pi'\}.$$

Si bien la notación $[\_]_\Sigma$ está sobrecargada, el argumento siempre desambiguará su uso.

Definición 7.2. Sean $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \{\mathrm{U}(i)\}_{i\in\mathsf{Agt}}, \mathrm{V}\rangle$ un LTS^U y Σ un conjunto de fórmulas de $\mathsf{L}_{\mathsf{Kh}_i}$ cerrado bajo subfórmulas. Para $i \in \mathsf{Agt}$ se define Act_i^Σ tal que $a_{[\pi]_\Sigma} \in \mathsf{Act}_i^\Sigma$ si y sólo si $\pi \in \mathrm{U}(i)$ es un testigo para algún $\mathsf{Kh}_i(\psi, \varphi)$ en $\mathfrak{M}$; y $\mathsf{Act}^\Sigma := \bigcup_{i\in\mathsf{Agt}} \mathsf{Act}_i^\Sigma$.

La idea de la definición de Act^Σ es la de definir una única acción que represente el comportamiento todos los conjuntos de planes π que coinciden en ser exactamente los testigos de un conjunto no vacío de subfórmulas de la forma $\mathsf{Kh}_i(\psi, \varphi)$ en $\mathfrak{M}$. Así como se habla de un colapso de estados, se busca también un colapso de los conjuntos de planes a una acción específica. Con esto ya se puede dar una noción de lo que es una filtración.

Definición 7.3 (Filtración de $\mathfrak{M}$ vía Σ)**.** Sean $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \{\mathrm{U}(i)\}_{i\in\mathsf{Agt}}, \mathrm{V}\rangle$ un LTS^U definido sobre Act y Σ un conjunto de fórmulas de $\mathsf{L}_{\mathsf{Kh}_i}$ cerrado bajo subfórmulas. Un LTS^U $\mathfrak{M}^f = \langle \mathrm{S}^f, \mathrm{R}^f, \{\mathrm{U}^f(i)\}_{i\in\mathsf{Agt}}, \mathrm{V}^f\rangle$ definido sobre $(\mathsf{Act}^\Sigma \cup \{a_\emptyset\})$ es una *filtración de* $\mathfrak{M}$ *vía* Σ si y sólo si cumple las siguientes condiciones:

1. $\mathrm{S}^f := \{[w]_\Sigma \mid w \in \mathrm{S}\}$;
2. $\mathrm{V}^f([w]_\Sigma) := \{p \in \Sigma \mid \mathfrak{M}, w \models p\}$;
3. $\mathrm{U}(i)^f := \{\{a_{[\pi]_\Sigma}\} \mid a_{[\pi]_\Sigma} \in \mathsf{Act}_i^\Sigma\} \cup \{\{a_\emptyset\}\}$ para cada $i \in \mathsf{Agt}$; y
4. $\mathrm{R}^f_{a_\emptyset} = \emptyset$, y para todo $a_{[\pi]_\Sigma} \in \mathsf{Act}_i^\Sigma$, existe un subconjunto no vacío $G_{[\pi]_\Sigma} \subseteq [\pi]_\Sigma$ tal que $([w]_\Sigma, [v]_\Sigma) \in \mathrm{R}^f_{a_{[\pi]_\Sigma}}$ sii
   - para todo $w' \in [w]_\Sigma$ y $\pi' \in G_{[\pi]_\Sigma}$, se tiene que $w' \in \mathrm{SE}(\pi')$, y
   - existen $w'' \in [w]_\Sigma$, $v'' \in [v]_\Sigma$ y $\pi'' \in G_{[\pi]_\Sigma}$ tales que $(w', v') \in \mathrm{R}_{\pi''}$;

La Definición 7.3 requiere algunas aclaraciones. En primer lugar, la misma es diferente a la vista en [10]. Esto se hace con el objetivo de tener filtraciones más manejables y convenientes en la práctica, dando condiciones más simples y fáciles de verificar. Volviendo a la definición actual, para la parte LTS, la filtración se define de manera similar a la que se hace en la lógica modal básica para S^f y V^f [24]. $\mathrm{R}^f_{a_{[\pi]_\Sigma}}$ se define con respecto a un subconjunto $G_{[\pi]_\Sigma} \subseteq [\pi]_\Sigma$, cuyos elementos coinciden en ser testigos de las mismas fórmulas Kh_i en Σ (Ítem 4). Para una clase de equivalencia $[w]_\Sigma$, se añaden arcos sólo si todo $\pi' \in G_{[\pi]_\Sigma}$ es fuertemente ejecutable en cada estado $w' \in [w]_\Sigma$. Se tienen que $\mathrm{S}^f \neq \emptyset$ y V^f está bien definida: dado $p \in \Sigma$, si $[w]_\Sigma = [v]_\Sigma$ y $\mathfrak{M}, w \models p$, entonces $\mathfrak{M}, v \models p$. Se puede ver que $\mathrm{U}^f(i)$ también está bien definida por construcción (Observación 4.1). Especialmente, de no existir ningún $\pi \in \mathrm{U}(i)$ tal que cumpla algún Kh_i en Σ, se sigue garantizando que $\mathrm{U}^f(i) \neq \emptyset$. Esto se debe a que se considera una acción vacua, $a_\emptyset$, que no es fuertemente ejecutable en ningún estado. Además, $(\mathsf{Act}^\Sigma \cup \{a_\emptyset\})$ es numerable. Por lo tanto, $\mathfrak{M}^f$ está bien definida. Finalmente, el Ítem 4 puede ser reescrito como $([w]_\Sigma, [v]_\Sigma) \in \mathrm{R}^f_{a_{[\pi]_\Sigma}}$ si y sólo si:

- $[w]_\Sigma \subseteq \bigcap_{\pi' \in G_{[\pi]_\Sigma}} \mathrm{SE}(\pi')$, y
- $(w'', v'') \in \bigcup_{\pi'' \in G_{[\pi]_\Sigma}} \mathrm{R}_{\pi''}$ para algún $w'' \in [w]_\Sigma$ y $v'' \in [v]_\Sigma$.

Más aún, $[w]_\Sigma \in \mathrm{SE}(a_{[\pi]_\Sigma})$ sii $[w]_\Sigma \subseteq \bigcap_{\pi' \in G_{[\pi]_\Sigma}} \mathrm{SE}(\pi')$. Con estas aclaraciones, procedemos a demostrar el siguiente teorema que, dado un modelo y su correspondiente filtración vía Σ, ambos satisfacen las mismas fórmulas de Σ.

Teorema 7.1. *Sean* $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \{\mathrm{U}(i)\}_{i\in\mathsf{Agt}}, \mathrm{V}\rangle$ *un* LTS^U *y* Σ *un conjunto de fórmulas de* $\mathsf{L}_{\mathsf{Kh}_i}$ *cerrado bajo subfórmulas. Entonces, para todo* $\psi \in \Sigma$ *y* $w \in \mathrm{S}$, $\mathfrak{M}, w \models \psi$ *sii* $\mathfrak{M}^f, [w]_\Sigma \models \psi$. *Más aún, si* Σ *es finito entonces* $\mathfrak{M}^f$ *es finito.*

Demostración. Los casos base y booleanos son bastante directos, ya sea usando la definición de $\models$ o la hipótesis inductiva. Por lo que se mostrará que si $\mathsf{Kh}_i(\psi,\varphi)\in\Sigma$, $\mathfrak{M},w\models\mathsf{Kh}_i(\psi,\varphi)$ sii $\mathfrak{M}^f,[w]_\Sigma\models\mathsf{Kh}_i(\psi,\varphi)$. Como $\mathsf{Kh}_i(\psi,\varphi)$ es una fórmula global, podemos omitir el punto de evaluación en los estados. Como consecuencia del Ítem 4, si $a_{[\pi]_\Sigma}\in\mathsf{Act}_i^\Sigma$, entonces $[w]_\Sigma\in\mathrm{SE}(a_{[\pi]_\Sigma})$ sii existen $w'\in[w]_\Sigma$ y $\pi'\in G_{[\pi]_\Sigma}$ tales que $w\in\mathrm{SE}(\pi')$.

($\Rightarrow$) Si $\mathfrak{M}\models\mathsf{Kh}_i(\psi,\varphi)$, sea $\pi\in\mathrm{U}(i)$ su correspondiente testigo. Por definición, $a_{[\pi]_\Sigma}\in\mathsf{Act}_i^\Sigma$ y por lo tanto, $\{a_{[\pi]_\Sigma}\}\in\mathrm{U}^f(i)$. Veamos que este es el testigo adecuado. Por el Ítem 4, para todo $\pi'\in G_{[\pi]_\Sigma}$, $\pi\leftrightarrows_\Sigma\pi'$. Luego, para todo $\pi'\in G_{[\pi]_\Sigma}$, $[\![\psi]\!]^{\mathfrak{M}}\subseteq\mathrm{SE}(\pi')$ y $\mathrm{R}_{\pi'}([\![\psi]\!]^{\mathfrak{M}})\subseteq[\![\varphi]\!]^{\mathfrak{M}}$.

1. Sea $[w]_\Sigma\in[\![\psi]\!]^{\mathfrak{M}^f}$. Por hipótesis inductiva, para cada $w'\in[w]_\Sigma=[w']_\Sigma$, $w'\in[\![\psi]\!]^{\mathfrak{M}}$. Con esto, $[w]_\Sigma\subseteq\mathrm{SE}(\pi')$ para todo $\pi'\in G_{[\pi]_\Sigma}$. Luego, $[w]_\Sigma\in\mathrm{SE}(a_{[\pi]_\Sigma})$. Por lo tanto, $[\![\psi]\!]^{\mathfrak{M}^f}\subseteq\mathrm{SE}(\{a_{[\pi]_\Sigma}\})$.

2. Sea $([w]_\Sigma,[v]_\Sigma)\in\mathrm{R}^f_{a_{[\pi]_\Sigma}}$ tal que $[w]_\Sigma\in[\![\psi]\!]^{\mathfrak{M}^f}$. Luego, existen $w'\in[w]_\Sigma=[w']_\Sigma$, $v'\in[v]_\Sigma=[v']_\Sigma$ y $\pi'\in G_{[\pi]_\Sigma}$ tales que $(w',v')\in\mathrm{R}_{\pi'}$. Por hipótesis inductiva, $w'\in[\![\psi]\!]^{\mathfrak{M}}$ y con esto $v'\in[\![\varphi]\!]^{\mathfrak{M}}$. Nuevamente por hipótesis inductiva, $[v']_\Sigma=[v]_\Sigma\in[\![\varphi]\!]^{\mathfrak{M}^f}$. Por lo tanto, $\mathrm{R}^f_{\{a_{[\pi]_\Sigma}\}}([\![\psi]\!]^{\mathfrak{M}^f})\subseteq[\![\varphi]\!]^{\mathfrak{M}^f}$.

De esta manera, $\mathfrak{M}^f\models\mathsf{Kh}_i(\psi,\varphi)$.

($\Leftarrow$) Si $\mathfrak{M}^f\models\mathsf{Kh}_i(\psi,\varphi)$, sea $\pi\in\mathrm{U}^f(i)$ su correspondiente testigo. Si $\pi=\{a_\emptyset\}$, entonces $[\![\psi]\!]^{\mathfrak{M}^f}\subseteq\mathrm{SE}(\{a_\emptyset\})=\emptyset$. Por hipótesis inductiva, $[\![\psi]\!]^{\mathfrak{M}}=\emptyset$ y cualquier $\pi\in\mathrm{U}(i)$ es testigo en $\mathfrak{M}$. Si $\pi=\{a_{[\pi']_\Sigma}\}$, se tiene que $[\![\psi]\!]^{\mathfrak{M}^f}\subseteq\mathrm{SE}(\{a_{[\pi']_\Sigma}\})$ y $\mathrm{R}^f_{\{a_{[\pi']_\Sigma}\}}([\![\psi]\!]^{\mathfrak{M}^f})\subseteq[\![\varphi]\!]^{\mathfrak{M}^f}$. Basta con demostrar que cualquier $\pi''\in G_{[\pi]_\Sigma}$ es el testigo adecuado.

1. Sea $w\in[\![\psi]\!]^{\mathfrak{M}}$, por hipótesis inductiva, $[w]_\Sigma\in[\![\psi]\!]^{\mathfrak{M}^f}$ y con esto $[w]_\Sigma\in\mathrm{SE}(\{a_{[\pi']_\Sigma}\})$. Por el Ítem 4, $[w]_\Sigma\subseteq\bigcap_{\pi''\in G_{[\pi']_\Sigma}}\mathrm{SE}(\pi'')$, y se tiene que $w\in\bigcap_{\pi''\in G_{[\pi']_\Sigma}}\mathrm{SE}(\pi'')$. Por lo tanto, $[\![\psi]\!]^{\mathfrak{M}}\subseteq\mathrm{SE}(\pi'')$, para todo $\pi''\in G_{[\pi']_\Sigma}$.

2. Sea $(w,v)\in\mathrm{R}_{\pi''}$ tal que $w\in[\![\psi]\!]^{\mathfrak{M}}$. Como $w\in\bigcap_{\pi''\in G_{[\pi']_\Sigma}}\mathrm{SE}(\pi'')$, necesariamente $([w]_\Sigma,[v]_\Sigma)\in\mathrm{R}^f_{a_{[\pi']_\Sigma}}$. Por hipótesis inductiva, $[w]_\Sigma\in[\![\psi]\!]^{\mathfrak{M}^f}$ y con esto $[v]_\Sigma\in[\![\varphi]\!]^{\mathfrak{M}^f}$. Nuevamente por hipótesis inductiva, $v\in[\![\varphi]\!]^{\mathfrak{M}}$. Por lo tanto, $\mathrm{R}_{\pi''}([\![\psi]\!]^{\mathfrak{M}})\subseteq[\![\varphi]\!]^{\mathfrak{M}}$ para todo $\pi''\in G_{[\pi']_\Sigma}$.

Luego, $\mathfrak{M}\models\mathsf{Kh}_i(\psi,\varphi)$. Queda demostrar que si Σ es finito, entonces $\mathfrak{M}^f$ también lo es. En primer lugar, el número de estados en S^f es a lo sumo 2^m, siendo m el número total de fórmulas en Σ. Por lo tanto, S^f y V^f son finitos. Por definición, para todo $i\in\mathsf{Agt}$, Act_i^Σ es a lo sumo exponencial en el número de $\mathsf{Kh}_i(\psi,\varphi)\in\Sigma$. Por lo tanto, $\mathsf{Act}^\Sigma\cup\{a_\emptyset\}$ también lo es. Con esto, R^f y U^f son finitos, y $\mathfrak{M}^f$ es finito. $\square$

Considerando el teorema anterior, basta entonces con demostrar que, dado un modelo $\mathfrak{M}$ y un conjunto de fórmulas Σ cerrado bajo subfórmulas, el conjunto de filtraciones de $\mathfrak{M}$ vía Σ es no vacío. Esto se puede ver definiendo $G_{[\pi]_\Sigma}=[\pi]_\Sigma\neq\emptyset$ para cada $a_{[\pi]_\Sigma}\in\mathsf{Act}_i^\Sigma$. Como consecuencia, el problema de satisfacibilidad para $\mathsf{L}_{\mathsf{Kh}_i}$ es decidible.

Corolario 7.1. *El problema de satisfacibilidad para* $\mathsf{L}_{\mathsf{Kh}_i}$ *es decidible.*

Demostración. Sea φ una fórmula de $\mathsf{L}_{\mathsf{Kh}_i}$. Si φ es satisfacible, entonces existe un LTS^U $\mathfrak{M}$ y $w\in\mathrm{D}_{\mathfrak{M}}$ tales que $\mathfrak{M},w\models\varphi$. Sea Σ el conjunto de todas las subfórmulas de φ (incluyendo la misma), se define el modelo $\mathfrak{M}^f=\langle\mathrm{S}^f,\mathrm{R}^f,\{\mathrm{U}^f(i)\}_{i\in\mathsf{Agt}},\mathrm{V}^f\rangle$ tal como está en la Definición 7.3, con $G_{[\pi]_\Sigma}=[\pi]_\Sigma$. Como $\mathfrak{M},w\models\varphi$, aplicando el Teorema 7.1, se tiene que $\mathfrak{M}^f,[w]_\Sigma\models\varphi$. Dado que Σ es finito, se tiene que $\mathfrak{M}^f$ es finito. Por lo tanto, para toda fórmula φ de $\mathsf{L}_{\mathsf{Kh}_i}$, la misma es satisfacible si y sólo si es satisfacible en un modelo finito. $\square$

En la próxima sección se refinará este resultado y caracterizaremos su complejidad exacta, tanto para su problema de satisfacibilidad como el de model checking. Para ello, se deberá analizar la construcción del modelo canónico de $\mathsf{L}_{\mathsf{Kh}_i}$, vista en el Capítulo 5, y realizar una selección de los elementos que la componen.

7.2. Complejidad vía selección

En esta sección se analiza la complejidad computacional del problema de satisfacibilidad de $\mathsf{L}_{\mathsf{Kh}_i}$ bajo la semántica basada en LTS^U. Se establecerá que dicho problema es NP-completo mostrando que, si una fórmula es satisfacible, entonces existe un modelo de tamaño polinomial que la satisface. Además, introduciremos un algoritmo de model checking que corre en tiempo polinomial.

Dada una fórmula, se mostrará que es posible seleccionar un fragmento del modelo canónico relevante para su evaluación. Este modelo seleccionado bastará para preservar la satisfacibilidad de la fórmula. Es más, el tamaño de este modelo será polinomial al tamaño de la fórmula dada.

Definición 7.4 (Función de selección)**.** Sea $\mathfrak{M}^\Gamma = \langle \mathrm{S}^\Gamma, \mathrm{R}^\Gamma, \{\mathrm{U}(i)^\Gamma\}_{i\in\mathsf{Agt}}, \mathrm{V}^\Gamma\rangle$ un modelo canónico para un conjunto maximal consistente Γ (Definición 5.3). Sean $w \in \mathrm{S}^\Gamma$ y una fórmula φ de $\mathsf{L}_{\mathsf{Kh}_i}$. Se define $\mathsf{Act}_\varphi := \{\langle\theta_1,\theta_2\rangle \in \mathsf{Act}^\Gamma \mid \mathsf{Kh}_i(\theta_1,\theta_2)$ es subfórmula de $\varphi\}$. Una *función de selección canónica* sel_w^φ es una función que toma un modelo $\mathfrak{M}^\Gamma$, $w \in \mathrm{S}^\Gamma$ y una fórmula φ, y devuelve un conjunto de estados $\mathrm{S}' \subseteq \mathrm{S}^\Gamma$ de la siguiente manera:

1. $\mathsf{sel}_w^\varphi(p) = \{w\}$;
2. $\mathsf{sel}_w^\varphi(\neg\varphi_1) = \mathsf{sel}_w^\varphi(\varphi_1)$;
3. $\mathsf{sel}_w^\varphi(\varphi_1 \vee \varphi_2) = \mathsf{sel}_w^\varphi(\varphi_1) \cup \mathsf{sel}_w^\varphi(\varphi_2)$;
4. Si $[\![\mathsf{Kh}_i(\varphi_1,\varphi_2)]\!]^{\mathfrak{M}^\Gamma} \neq \emptyset$ y $[\![\varphi_1]\!]^{\mathfrak{M}^\Gamma} = \emptyset$:
   $\mathsf{sel}_w^\varphi(\mathsf{Kh}_i(\varphi_1,\varphi_2)) = \{w\}$;
5. Si $[\![\mathsf{Kh}_i(\varphi_1,\varphi_2)]\!]^{\mathfrak{M}^\Gamma} \neq \emptyset$ y $[\![\varphi_1]\!]^{\mathfrak{M}^\Gamma} \neq \emptyset$:
   $\mathsf{sel}_w^\varphi(\mathsf{Kh}_i(\varphi_1,\varphi_2)) = \{w_1,w_2\} \cup \mathsf{sel}_{w_1}^\varphi(\varphi_1) \cup \mathsf{sel}_{w_2}^\varphi(\varphi_2)$,
   donde w_1, w_2 son tal que $(w_1,w_2) \in \mathrm{R}^\Gamma_{\langle\varphi_1,\varphi_2\rangle}$;
6. Si $[\![\mathsf{Kh}_i(\varphi_1,\varphi_2)]\!]^{\mathfrak{M}^\Gamma} = \emptyset$ (notar que $[\![\varphi_1]\!]^{\mathfrak{M}^\Gamma} \neq \emptyset$):
   Para cada $\pi \in \mathrm{U}(i)^\Gamma$, o bien $[\![\varphi_1]\!]^{\mathfrak{M}^\Gamma} \not\subseteq \mathrm{SE}(\pi)$, o $\mathrm{R}^\Gamma_\pi([\![\varphi_1]\!]^{\mathfrak{M}^\Gamma}) \not\subseteq [\![\varphi_2]\!]^{\mathfrak{M}^\Gamma}$. Para cada $a \in \mathsf{Act}_\varphi$:
   a) si $[\![\varphi_1]\!]^{\mathfrak{M}^\Gamma} \not\subseteq \mathrm{SE}(\{a\})$,
   se agrega $\{w_1\} \cup \mathsf{sel}_{w_1}^\varphi(\varphi_1)$ a $\mathsf{sel}_w^\varphi(\mathsf{Kh}_i(\varphi_1,\varphi_2))$,
   donde $w_1 \in [\![\varphi_1]\!]^{\mathfrak{M}^\Gamma}$ y $w_1 \notin \mathrm{SE}(\{a\})$;
   b) si $[\![\varphi_1]\!]^{\mathfrak{M}^\Gamma} \subseteq \mathrm{SE}(\{a\})$ pero $\mathrm{R}^\Gamma_\pi([\![\varphi_1]\!]^{\mathfrak{M}^\Gamma}) \not\subseteq [\![\varphi_2]\!]^{\mathfrak{M}^\Gamma}$,
   se agrega $\{w_1,w_2\} \cup \mathsf{sel}_{w_1}^\varphi(\varphi_1) \cup \mathsf{sel}_{w_2}^\varphi(\varphi_2)$ a $\mathsf{sel}_w^\varphi(\mathsf{Kh}_i(\varphi_1,\varphi_2))$,
   donde $w_1 \in [\![\varphi_1]\!]^{\mathfrak{M}^\Gamma}$, $w_2 \in \mathrm{R}^\Gamma_a(w_1)$ y $w_2 \notin [\![\varphi_2]\!]^{\mathfrak{M}^\Gamma}$.

Con esto, dada una fórmula satisfacible, se puede seleccionar un modelo reducido que preserva su satisfacibilidad.

Definición 7.5 (Modelo seleccionado)**.** Sea $\mathfrak{M}^\Gamma$ un modelo canónico para un conjunto maximal consistente Γ, w un estado de $\mathfrak{M}^\Gamma$, y φ una fórmula de $\mathsf{L}_{\mathsf{Kh}_i}$. Sea sel_w^φ una función de selección, se define el *modelo seleccionado* por sel_w^φ como $\mathfrak{M}_w^\varphi = \langle \mathrm{S}_w^\varphi, \mathrm{R}_w^\varphi, \{(\mathrm{U}_w^\varphi)(i)\}_{i\in\mathsf{Agt}}, \mathrm{V}_w^\varphi\rangle$, donde

- $\mathrm{S}_w^\varphi = \mathsf{sel}_w^\varphi(\varphi)$;
- $(\mathrm{R}_w^\varphi)_{\langle\theta_1,\theta_2\rangle} = \mathrm{R}^\Gamma_{\langle\theta_1,\theta_2\rangle} \cap (\mathrm{S}_w^\varphi \times \mathrm{S}_w^\varphi)$, para cada $\langle\theta_1,\theta_2\rangle \in \mathsf{Act}_\varphi$;
- $(\mathrm{U}_w^\varphi)(i) = \{\{a\} \mid a \in \mathsf{Act}_\varphi\} \cup \{\{\langle\bot,\top\rangle\}\}$, para $i \in \mathsf{Agt}$ (y $(\mathrm{R}_w^\varphi)_{\langle\bot,\top\rangle} = \emptyset$);
- V_w^φ es una restricción de V^Γ a S_w^φ.

Notar que, a pesar de que Act_φ puede ser un conjunto vacío, cada colección de conjuntos de planes $(\mathrm{U}_w^\varphi)(i)$ no lo es. Más aún, Act_φ puede ser extendido a un conjunto infinito de acciones, para ser definido sobre una signatura adecuada.

Proposición 7.1. *$\mathfrak{M}^{\varphi}_{w} = \langle \mathrm{S}^{\varphi}_{w}, \mathrm{R}^{\varphi}_{w}, \mathrm{V}^{\varphi}_{w}, \mathrm{U}^{\varphi}_{w} \rangle$ es un* LTS^{U}.

Demostración. La estructura $\langle \mathrm{S}^{\varphi}_{w}, \mathrm{R}^{\varphi}_{w}, \mathrm{V}^{\varphi}_{w} \rangle$ es un LTS dado que $\mathsf{sel}^{\varphi}_{w}(\varphi) \neq \emptyset$. Como $(\mathrm{U}^{\varphi}_{w})$ es una restricción de U^{Γ} a $\mathsf{Act}_{\varphi} \cup \{\langle \bot, \top \rangle\}$, es fácil ver que se cumplen las condiciones de Definición 4.2. Por lo tanto, $\mathfrak{M}^{\varphi}_{w}$ es un LTS^{U}. $\square$

Con esto procedemos a demostrar que este modelo seleccionado preserva la satisfacibilidad de la fórmula en el modelo canónico.

Proposición 7.2. *Sea $\mathfrak{M}^{\Gamma}$ un modelo canónico, w un estado de $\mathfrak{M}^{\Gamma}$ y φ una fórmula de $\mathsf{L}_{\mathsf{Kh}_i}$. Sea $\mathfrak{M}^{\varphi}_{w}$ el modelo seleccionado por una función de selección $\mathsf{sel}^{\varphi}_{w}$. Entonces, $\mathfrak{M}^{\Gamma}, w \models \varphi$ implica que para toda ψ subfórmula de φ, y para toda $v \in \mathrm{S}^{\varphi}_{w}$ se tiene que $\mathfrak{M}^{\Gamma}, v \models \psi$ si y sólo si $\mathfrak{M}^{\varphi}_{w}, v \models \psi$. Más aún, $\mathfrak{M}^{\varphi}_{w}$ es de tamaño polinomial al tamaño de la fórmula φ.*

Demostración. La prueba se hace por inducción en el tamaño de la fórmula:

- Caso $\psi = p$: si $\mathfrak{M}^{\Gamma}, v \models p$, entonces $p \in \mathrm{V}^{\Gamma}(v)$. Dado que $v \in \mathrm{S}^{\varphi}_{w}$, se tiene que $p \in \mathrm{V}^{\varphi}_{w}(v)$ y por lo tanto $\mathfrak{M}^{\varphi}_{w}, v \models p$. La otra dirección es similar.
- Caso $\psi = \neg\psi_1$: si $\mathfrak{M}^{\Gamma}, w \models \neg\psi_1$, entonces $\mathfrak{M}^{\Gamma}, w \not\models \psi_1$. Por hipótesis inductiva, $\mathfrak{M}^{\varphi}_{w}, w \not\models \psi_1$ y por lo tanto $\mathfrak{M}^{\varphi}_{w}, w \models \neg\psi_1$. La otra dirección es similar.
- Caso $\psi = \psi_1 \vee \psi_2$: si $\mathfrak{M}^{\Gamma}, v \models \psi_1 \vee \psi_2$, entonces $\mathfrak{M}^{\Gamma}, v \models \psi_1$ ó $\mathfrak{M}^{\Gamma}, v \models \psi_2$. Por hipótesis inductiva, $\mathfrak{M}^{\varphi}_{w}, v \models \psi_1$ ó $\mathfrak{M}^{\varphi}_{w}, v \models \psi_2$ y por lo tanto $\mathfrak{M}^{\varphi}_{w}, v \models \psi_1 \vee \psi_2$. La otra dirección es similar.
- Caso $\psi = \mathsf{Kh}_i(\psi_1, \psi_2)$: Si $\mathfrak{M}^{\Gamma}, v \models \mathsf{Kh}_i(\psi_1, \psi_2)$, hay dos posibilidades:
  - $[\![\psi_1]\!]^{\mathfrak{M}^{\Gamma}} = \emptyset$: Como $\mathfrak{M}^{\Gamma}, v \models \mathsf{Kh}_i(\psi_1, \psi_2)$ existe $\pi \in \mathrm{U}^{\Gamma}(i)$ tal que $\emptyset = [\![\psi_1]\!]^{\mathfrak{M}^{\Gamma}} \subseteq \mathrm{SE}^{\mathfrak{M}^{\Gamma}}(\pi)$ y $\emptyset = \mathrm{R}^{\Gamma}_{\pi}([\![\psi_1]\!]^{\mathfrak{M}^{\Gamma}}) \subseteq [\![\psi_2]\!]^{\mathfrak{M}^{\Gamma}}$. Por hipótesis inductiva, $[\![\psi_1]\!]^{\mathfrak{M}^{\varphi}_{w}} \subseteq [\![\psi_1]\!]^{\mathfrak{M}^{\Gamma}}$. Notar que, como $[\![\psi_1]\!]^{\mathfrak{M}^{\Gamma}} = \emptyset$, se tiene que $[\![\psi_1]\!]^{\mathfrak{M}^{\varphi}_{w}} = \emptyset$. Sea $\pi' = \{\langle \bot, \top \rangle\}$, $\pi' \in (\mathrm{U}^{\varphi}_{w})(i)$, y $(\mathrm{R}^{\varphi}_{w})_{\pi'}([\![\psi_1]\!]^{\mathfrak{M}^{\varphi}_{w}}) = \emptyset$. Luego, existe $\pi' \in (\mathrm{U}^{\varphi}_{w})(i)$ tal que $[\![\psi_1]\!]^{\mathfrak{M}^{\varphi}_{w}} \subseteq \mathrm{SE}^{\mathfrak{M}^{\varphi}_{w}}(\pi')$ y $(\mathrm{R}^{\varphi}_{w})_{\pi'}([\![\psi_1]\!]^{\mathfrak{M}^{\varphi}_{w}}) \subseteq [\![\psi_2]\!]^{\mathfrak{M}}$. Por lo tanto, $\mathfrak{M}^{\varphi}_{w}, v \models \mathsf{Kh}_i(\psi_1, \psi_2)$.
  - $[\![\psi_1]\!]^{\mathfrak{M}^{\Gamma}} \neq \emptyset$: Como $\mathfrak{M}^{\Gamma}, v \models \mathsf{Kh}_i(\psi_1, \psi_2)$ existe $\pi \in \mathrm{U}^{\Gamma}(i)$ tal que $[\![\psi_1]\!]^{\mathfrak{M}^{\Gamma}} \subseteq \mathrm{SE}^{\mathfrak{M}^{\Gamma}}(\pi)$ y $\mathrm{R}^{\Gamma}_{\pi}([\![\psi_1]\!]^{\mathfrak{M}^{\Gamma}}) \subseteq [\![\psi_2]\!]^{\mathfrak{M}^{\Gamma}}$. Usando el Lema 5.1, $\mathsf{Kh}_i(\psi_1, \psi_2) \in v$, y aplicando la definición del modelo canónico, $\mathsf{Kh}_i(\psi_1, \psi_2) \in \Gamma$ y $\langle \psi_1, \psi_2 \rangle \in \mathsf{Act}_{\Gamma}$. Por la definición de $\mathrm{R}^{\Gamma}_{\langle \psi_1, \psi_2 \rangle}$, se tiene que para todo $w \in [\![\psi_1]\!]^{\mathfrak{M}^{\Gamma}}$, se cumple que $\mathrm{R}^{\Gamma}_{\langle \psi_1, \psi_2 \rangle}(w) \neq \emptyset$ y $\mathrm{R}^{\Gamma}_{\langle \psi_1, \psi_2 \rangle}(w) \subseteq [\![\psi_2]\!]^{\mathfrak{M}^{\Gamma}}$. Por lo tanto, $[\![\psi_1]\!]^{\mathfrak{M}^{\Gamma}} \subseteq \mathrm{SE}^{\mathfrak{M}^{\Gamma}}(\langle \psi_1, \psi_2 \rangle)$ y $\mathrm{R}^{\Gamma}_{\langle \psi_1, \psi_2 \rangle}([\![\psi_1]\!]^{\mathfrak{M}^{\Gamma}}) \subseteq [\![\psi_2]\!]^{\mathfrak{M}^{\Gamma}}$. Como $[\![\psi_1]\!]^{\mathfrak{M}^{\Gamma}} \neq \emptyset$, existen $w_1, w_2 \in \mathrm{S}^{\Gamma}$ tales que $(w_1, w_2) \in \mathrm{R}^{\Gamma}_{\langle \psi_1, \psi_2 \rangle}$.
    Por definición de $\mathfrak{M}^{\varphi}_{w}$, se tiene que $\{\langle \psi_1, \psi_2 \rangle\} \in (\mathrm{U}^{\varphi}_{w})(i)$ y que $(\mathrm{R}^{\varphi}_{w})_{\langle \psi_1, \psi_2 \rangle}$ está definida. Además, por definición de $\mathsf{sel}^{\varphi}_{w}$, Ítem 5, existen $w'_1, w'_2 \in \mathrm{S}^{\varphi}_{w}$ tales que $(w'_1, w'_2) \in (\mathrm{R}^{\varphi}_{w})_{\langle \psi_1, \psi_2 \rangle}$. Sea $v_1 \in [\![\psi_1]\!]^{\mathfrak{M}^{\varphi}_{w}}$, por hipótesis inductiva $v_1 \in [\![\psi_1]\!]^{\mathfrak{M}^{\Gamma}}$. Luego se tiene que $v_1 \in \mathrm{SE}^{\mathfrak{M}^{\Gamma}}(\{\langle \psi_1, \psi_2 \rangle\})$ y $\mathrm{R}^{\Gamma}_{\langle \psi_1, \psi_2 \rangle}(v_1) \subseteq [\![\psi_2]\!]^{\mathfrak{M}^{\Gamma}}$. Como para todo $v_2 \in \mathrm{R}^{\Gamma}_{\langle \psi_1, \psi_2 \rangle}(v_1)$, se tiene que $v_2 \in [\![\psi_2]\!]^{\mathfrak{M}^{\Gamma}}$, particularmente $v_2 = w'_2$, entonces $w'_2 \in (\mathrm{R}^{\varphi}_{w})_{\langle \psi_1, \psi_2 \rangle}(v_1)$. Luego, $v_1 \in \mathrm{SE}^{\mathfrak{M}^{\varphi}_{w}}(\{\langle \psi_1, \psi_2 \rangle\})$.
    Buscando una contradicción, si $(\mathrm{R}^{\varphi}_{w})_{\langle \psi_1, \psi_2 \rangle}(v_1) = \mathrm{R}^{\Gamma}_{\langle \psi_1, \psi_2 \rangle}(v_1) \cap \mathrm{S}^{\varphi}_{w} \not\subseteq [\![\psi_2]\!]^{\mathfrak{M}^{\varphi}_{w}}$, entonces existe $v_2 \in (\mathrm{R}^{\varphi}_{w})_{\langle \psi_1, \psi_2 \rangle}(v_1)$ tal que $v_2 \notin [\![\psi_2]\!]^{\mathfrak{M}^{\varphi}_{w}}$. Luego $(\mathrm{R}^{\varphi}_{w})_{\langle \psi_1, \psi_2 \rangle}(v_1) \subseteq \mathrm{R}^{\Gamma}_{\langle \psi_1, \psi_2 \rangle}(v_1)$, pero por hipótesis inductiva $v_2 \notin [\![\psi_2]\!]^{\mathfrak{M}^{\Gamma}}$. Por lo tanto, $\{\langle \psi_1, \psi_2 \rangle\}$ no es testigo de $\mathsf{Kh}_i(\psi_1, \psi_2)$ en $\mathfrak{M}^{\Gamma}$, lo cual es una contradicción. Luego, $(\mathrm{R}^{\varphi}_{w})_{\{\langle \psi_1, \psi_2 \rangle\}}(v_1) \subseteq [\![\psi_2]\!]^{\mathfrak{M}^{\varphi}_{w}}$. Dado que $[\![\psi_1]\!]^{\mathfrak{M}^{\varphi}_{w}} \subseteq \mathrm{SE}^{\mathfrak{M}^{\varphi}_{w}}(\{\langle \psi_1, \psi_2 \rangle\})$ y $(\mathrm{R}^{\varphi}_{w})_{\{\langle \psi_1, \psi_2 \rangle\}}([\![\psi_1]\!]^{\mathfrak{M}^{\varphi}_{w}}) \subseteq [\![\psi_2]\!]^{\mathfrak{M}^{\varphi}_{w}}$, se tiene que $\mathfrak{M}^{\varphi}_{w}, v \models \mathsf{Kh}_i(\psi_1, \psi_2)$.
- Si $\mathfrak{M}^{\varphi}_{w}, v \models \mathsf{Kh}_i(\psi_1, \psi_2)$, hay dos posibilidades:
  - $[\![\psi_1]\!]^{\mathfrak{M}^{\varphi}_{w}} = \emptyset$: Como $\mathfrak{M}^{\varphi}_{w}, v \models \mathsf{Kh}_i(\psi_1, \psi_2)$, entonces $\emptyset = [\![\psi_1]\!]^{\mathfrak{M}^{\varphi}_{w}} \subseteq \mathrm{SE}^{\mathfrak{M}^{\varphi}_{w}}(\pi')$ y $\emptyset = (\mathrm{R}^{\varphi}_{w})_{\pi'}([\![\psi_1]\!]^{\mathfrak{M}^{\varphi}_{w}}) \subseteq [\![\psi_2]\!]^{\mathfrak{M}^{\varphi}_{w}}$ para algún $\pi' \in (\mathrm{U}^{\varphi}_{w})(i)$. Necesariamente $[\![\psi_1]\!]^{\mathfrak{M}^{\Gamma}} = \emptyset$. Si $\mathfrak{M}^{\Gamma}, v \models \mathsf{Kh}_i(\psi_1, \psi_2)$, por $\mathsf{sel}^{\varphi}_{w}$, Ítem 5, $\emptyset \neq (\mathrm{R}^{\varphi}_{w})_{\langle \psi_1, \psi_2 \rangle}$ está definida y $[\![\psi_1]\!]^{\mathfrak{M}^{\varphi}_{w}} \neq \emptyset$,

contradiciendo la hipótesis. Y si $\mathfrak{M}^\Gamma, v \not\models \mathsf{Kh}_i(\psi_1, \psi_2)$, por sel^φ_w, Ítem 6 e hipótesis inductiva, $[\![\psi_1]\!]^{\mathfrak{M}^\varphi_w} \neq \emptyset$, que también es una contradicción.

Sea π cualquier conjunto de planes en $\mathrm{U}^\Gamma(i)$. Como $\mathrm{R}^\Gamma_\pi([\![\psi_1]\!]^{\mathfrak{M}^\Gamma}) = \emptyset$, $[\![\psi_1]\!]^{\mathfrak{M}^\Gamma} \subseteq \mathrm{SE}(\pi)$ y $\mathrm{R}^\Gamma_\pi([\![\psi_1]\!]^{\mathfrak{M}^\Gamma}) \subseteq [\![\psi_2]\!]^{\mathfrak{M}^\Gamma}$. Luego, $\mathfrak{M}^\Gamma, v \models \mathsf{Kh}_i(\psi_1, \psi_2)$.

- $[\![\psi_1]\!]^{\mathfrak{M}^\varphi_w} \neq \emptyset$: Notar que por hipótesis inductiva $[\![\psi_1]\!]^{\mathfrak{M}^\Gamma} \neq \emptyset$. Además, por $\mathfrak{M}^\varphi_w, v \models \mathsf{Kh}_i(\psi_1, \psi_2)$, se tiene que $[\![\psi_1]\!]^{\mathfrak{M}^\varphi_w} \subseteq \mathrm{SE}^{\mathfrak{M}^\varphi_w}(\pi')$ y $(\mathrm{R}^\varphi_w)_{\pi'}([\![\psi_1]\!]^{\mathfrak{M}^\varphi_w}) \subseteq [\![\psi_2]\!]^{\mathfrak{M}^\varphi_w}$, para algún $\pi' \in (\mathrm{U}^\varphi_w)(i)$. Buscando una contradicción, si $\mathfrak{M}^\Gamma, v \not\models \mathsf{Kh}_i(\psi_1, \psi_2)$, entonces para todo $\pi \in \mathrm{U}^\Gamma(i)$, $[\![\psi_1]\!]^{\mathfrak{M}^\Gamma} \not\subseteq \mathrm{SE}^{\mathfrak{M}^\Gamma}(\pi)$ ó $\mathrm{R}^\Gamma_\pi([\![\psi_1]\!]^{\mathfrak{M}^\Gamma}) \not\subseteq [\![\psi_2]\!]^{\mathfrak{M}^\Gamma}$. Por la definición de Act_φ, para todo $\pi = \{a\} \in (\mathrm{U}^\varphi_w)(i)$, con $a \in \mathsf{Act}_\varphi$, $[\![\psi_1]\!]^{\mathfrak{M}^\Gamma} \not\subseteq \mathrm{SE}^{\mathfrak{M}^\Gamma}(\pi)$ ó $\mathrm{R}^\Gamma_\pi([\![\psi_1]\!]^{\mathfrak{M}^\Gamma}) \not\subseteq [\![\psi_2]\!]^{\mathfrak{M}^\Gamma}$. Es decir, para todo $a \in \mathsf{Act}_\varphi$, $[\![\psi_1]\!]^{\mathfrak{M}^\Gamma} \not\subseteq \mathrm{SE}^{\mathfrak{M}^\Gamma}(\{a\})$ ó $\mathrm{R}^\Gamma_{\{a\}}([\![\psi_1]\!]^{\mathfrak{M}^\Gamma}) \not\subseteq [\![\psi_2]\!]^{\mathfrak{M}^\Gamma}$. Luego existe $w_1 \in [\![\psi_1]\!]^{\mathfrak{M}^\Gamma}$ tal que $w_1 \notin \mathrm{SE}^{\mathfrak{M}^\Gamma}(a)$ o existe $w_2 \in \mathrm{R}^\Gamma_a(w_1)$ tal que $w_2 \notin [\![\psi_2]\!]^{\mathfrak{M}^\Gamma}$. Por definición de sel^φ_w, Ítem 6, se agregan los estados testigo para cada $a \in \mathsf{Act}_\varphi$. Sea $\pi' \in (\mathrm{U}^\varphi_w)(i)$. Si $\pi' = \{\langle\bot, \top\rangle\}$, se obtiene trivialmente que $\emptyset \neq [\![\psi_1]\!]^{\mathfrak{M}^\varphi_w} \not\subseteq \mathrm{SE}^{\mathfrak{M}^\varphi_w}(\pi') = \emptyset$. Tomando un $\pi' = \{a\}$ tal que $a \in \mathsf{Act}_\varphi$, y el estado seleccionado $w'_1 \in [\![\psi_1]\!]^{\mathfrak{M}^\varphi_w} \subseteq [\![\psi_1]\!]^{\mathfrak{M}^\Gamma}$. Si $w'_1 \notin \mathrm{SE}^{\mathfrak{M}^\Gamma}(a)$, $\mathrm{R}^\Gamma_a(w'_1) = \emptyset$, entonces $(\mathrm{R}^\varphi_w)_a(w'_1) = \emptyset$ y por lo tanto $w'_1 \notin \mathrm{SE}^{\mathfrak{M}^\varphi_w}(\{a\})$. Por otro lado, si existe $w_2 \in \mathrm{R}^\Gamma_a(w'_1)$ tal que $w_2 \notin [\![\psi_2]\!]^{\mathfrak{M}^\Gamma}$, entonces por sel^φ_w e hipótesis inductiva, existe $w'_2 \in \mathrm{S}^\varphi_w$ estado seleccionado tal que $w'_2 \in \mathrm{R}^\Gamma_a(w'_1)$ y $w'_2 \notin [\![\psi_2]\!]^{\mathfrak{M}^\varphi_w}$. Con esto, existe $w'_2 \in (\mathrm{R}^\varphi_w)_a(w'_1)$ tal que $w'_2 \notin [\![\psi_2]\!]^{\mathfrak{M}^\varphi_w}$. En cualquier caso, esto implica que $\mathfrak{M}^\varphi_w, v \not\models \mathsf{Kh}_i(\psi_1, \psi_2)$, una contradicción. Por lo tanto, $\mathfrak{M}^\Gamma, v \models \mathsf{Kh}_i(\psi_1, \psi_2)$.

Dado que el modelo se construye a partir de $\mathsf{sel}^\varphi_w(\varphi)$, la función de selección agrega estados de $\mathfrak{M}^\Gamma$ únicamente cuando se considera una subfórmula de la forma $\mathsf{Kh}_i(\varphi_1, \varphi_1)$. Por lo tanto, $\mathsf{sel}^\varphi_w(\varphi)$ evalúa el caso Kh_i una cantidad polinomial en el tamaño de φ. Además, como el número de estados que se agregan es polinomial en el tamaño de φ, S^φ_w también lo es. Con un razonamiento análogo, Act_φ y $(\mathrm{U}^\varphi_w)(i)$ también son polinomiales. Con esto, $\mathfrak{M}^\varphi_w$ tiene tamaño polinomial.

Finalmente, la definición de un LTS^U requiere que el conjunto de acciones sea infinito. Es fácil ver que Act_φ se puede extender a un conjunto infinito. Como la Definición 3.4 establece relaciones sobre un subconjunto de acciones, esta extensión no altera el tamaño de $\mathfrak{M}^\varphi_w$. □

Para demostrar que el problema de satisfacibilidad de $\mathsf{L}_{\mathsf{Kh}_i}$ está en NP, necesitamos mostrar que el problema de model checking está en P.

Proposición 7.3. *El problema de model checking para* $\mathsf{L}_{\mathsf{Kh}_i}$ *está en* P.

Demostración. Dado un LTS^U $\mathfrak{M}$, un estado $w \in \mathrm{D}_{\mathfrak{M}}$ y una fórmula φ, se define un algoritmo de etiquetado de abajo hacia arriba (bottom-up) en tiempo polinomial que verifica si $\mathfrak{M}, w \models \varphi$. Se siguen las mismas ideas de la lógica modal básica ML [26]. A continuación se introduce el caso para las fórmulas de la forma $\mathsf{Kh}_i(\psi, \varphi)$, sobre un LTS^U $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \mathrm{U}, \mathrm{V}\rangle$, $w \in \mathrm{S}$:

Algoritmo 1: Model Checking para Kh_i

```
Datos: 𝔐 un LTS^U, w ∈ D_𝔐, Kh_i(ψ, φ)
Resultado: kh
lab(Kh_i(ψ, φ)) ← ∅;
para todo π ∈ U(i) hacer
    kh ← Verdadero;
    para todo σ ∈ π hacer
        para todo v ∈ lab(ψ) hacer
            kh ← (kh & v ∈ SE(σ) & R_σ(v) ⊆ lab(φ));
        fin
    fin
    si kh entonces
        lab(Kh_i(ψ, φ)) ← S;
    fin
fin
```

Como $\mathrm{U}(i)$ y cada $\pi \in \mathrm{U}(i)$ son no vacíos, los primeros dos ciclos **para todo** se ejecutan necesariamente. Si $lab(\psi) = \emptyset$, entonces la fórmula $\mathsf{Kh}_i(\psi, \varphi)$ es trivialmente verdadera. Si

$lab(\psi) \neq \emptyset$, kh se mantendrá verdadera sólo si las condiciones apropiadas para la satisfacibilidad de $\mathsf{Kh}_i(\psi, \varphi)$ se cumplen para algún $\pi \in \mathrm{U}(i)$. Caso contrario, la inicialización de $lab(\mathsf{Kh}_i(\psi, \varphi))$ como $\emptyset$ no se sobrescribe. Ambos $v \in \mathrm{SE}(\sigma)$ y $\mathrm{R}_\sigma(v) \subseteq lab(\varphi)$ pueden ser verificados en tiempo polinomial. Por lo tanto, el problema de model checking está en P. $\square$

Con esto, se tiene el resultado para el problema de satisfacibilidad.

Teorema 7.2. *El problema de satisfacibilidad para $\mathsf{L}_{\mathsf{Kh}_i}$ sobre LTS^Us es NP-completo.*

Demostración. Hardness para NP (es decir, la cota inferior) se sigue de que la lógica proposicional, un fragmento de $\mathsf{L}_{\mathsf{Kh}_i}$, es NP-completo [12, 35]. Por la Proposición 7.2, cada fórmula satisfacible φ tiene un modelo de tamaño polinomial con respecto a φ. Con esto, se puede adivinar un modelo polinomial $\mathfrak{M}$ y un estado $w \in \mathrm{D}_{\mathfrak{M}}$, y verificar $\mathfrak{M}, w \models \varphi$ en tiempo polinomial, dado el resultado de la Proposición 7.3. Por lo tanto, el problema de satisfacibilidad está en NP. $\square$

Para finalizar esta sección, como corolario, se tiene que para toda fórmula φ de $\mathsf{L}_{\mathsf{Kh}_i}$, φ es satisfacible si y sólo si φ es satisfacible en un modelo finito, tal que los conjuntos de planes de cada $\mathrm{U}(i)$ son de un solo elemento de longitud uno. Esto es similar a la "tree model property" de la lógica modal básica, la cual establece la satisfacibilidad de una fórmula es equivalente a que sea satisfacible en un modelo en forma de árbol, siendo la raíz el nodo donde se evalúa la fórmula [24]. Este resultado es importante ya que nos servirá para estudiar extensiones de $\mathsf{L}_{\mathsf{Kh}_i}$ con operadores dinámicos en el Capítulo 8, especialmente para saber si el problema de satisfacibilidad es decidible.

Corolario 7.2. *Sea la siguiente la clase de todos los modelos LTS^Us finitos, los cuales sus agentes consideran únicamente acciones básicas distinguibles entre sí (**FNU**):*

$$\mathbf{M}^{\mathbf{F}}_{\mathbf{NU}} := \{\mathfrak{M} \mid \mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \mathrm{U}, \mathrm{V}\rangle \textit{ es un } \mathrm{LTS}^U \textit{ finito y para todo } i \in \mathsf{Agt},\ \mathrm{U}(i) \subseteq \{\{a\} \mid a \in \mathsf{Act}\}\}.$$

Se tiene que φ es satisfacible si y sólo si φ es satisfacible en $\mathbf{M}^{\mathbf{F}}_{\mathbf{NU}}$, es decir, existen $\mathfrak{M} \in \mathbf{M}^{\mathbf{F}}_{\mathbf{NU}}$ y $w \in \mathrm{D}_{\mathfrak{M}}$ tales que $\mathfrak{M}, w \models \varphi$.

Demostración. Dado el Teorema 5.1, $\{\varphi\}$ es satisfacible sii $\{\varphi\}$ es consistente. Utilizando la demostración de dicho teorema, si $\{\varphi\}$ es satisfacible, sea Γ una extensión maximal consistente de $\{\varphi\}$, se tiene que $\mathfrak{M}^\Gamma, \Gamma \models \varphi$ (Lema 5.1). Sea $\mathfrak{M}^\varphi_\Gamma$ el modelo seleccionado por la función de selección $\mathsf{sel}^\varphi_\Gamma$. Por la Proposición 7.2, se tiene que $\mathfrak{M}^\varphi_\Gamma, \Gamma \models \varphi$. $\mathfrak{M}^\varphi_\Gamma$ es de tamaño polinomial al tamaño de la fórmula φ y por lo tanto es finito. Por la Definición 7.5, $\mathrm{U}^\varphi_\Gamma$ cumple la segunda condición y por lo tanto, $\mathfrak{M}^\varphi_\Gamma \in \mathbf{M}^{\mathbf{F}}_{\mathbf{NU}}$. Con esto, si φ es satisfacible, entonces φ es satisfacible en $\mathbf{M}^{\mathbf{F}}_{\mathbf{NU}}$. La vuelta es directa. $\square$

Parte III

Explotando el framework

Capítulo 8

Lógicas dinámicas epistémicas de saber cómo

> But the Fox had so many plans for escape he could not decide which one to try first. He dodged here and there with the hounds at his heels. He doubled on his tracks, he ran at top speed, he entered a dozen burrows,—but all in vain. The hounds caught him, and soon put an end to the boaster and all his tricks.
>
> The Fox and the Cat

La nueva semántica basada en incertidumbre entre planes trae consigo varias ventajas. En particular, dicha relación de indistinguibilidad epistémica (representada por el conjunto $\mathrm{U}(i)$ para cada agente i) puede ser tratada de manera similar a la relación de indistinguibilidad entre mundos en lógica epistémica clásica [93, 108]. Por ejemplo, desde un punto de vista conceptual, permite una clara distinción entre la información óntica o factual (dada por la estructura del LTS) que es común a todos los agentes, y la información epistémica o percepción de cada agente (dada por el conjunto $\mathrm{U}(i)$). Por otro lado, es posible utilizar esta nueva componente para analizar otras nociones básicas como el conocimiento distribuido o colectivo (ver [39, 53] para más detalles).

Por otro lado, en lógica epistémica, usualmente es deseable representar el cambio de la información que cada agente posee mediante la ejecución de ciertas acciones. Dichas acciones en general se representan mediante operadores dinámicos dentro del lenguaje, dando origen a las lógicas dinámicas epistémicas [19, 109]. Un operador dinámico en este contexto es, en pocas palabras, un operador lógico que modifica el modelo al momento de ser evaluado (para una discusión más detallada, ver por ejemplo [7, 69, 91]). En lógica epistémica, esto se traduce en cambiar la información que percibe un agente, mediante modificaciones en los estados alcanzados, la relación entre los mismos y/o la valuación (ver por ejemplo [69, 91]). De la misma manera, en las lógicas de *saber cómo*, podemos definir operaciones que actualicen la información de un agente, tanto óntica como epistémica. El siguiente ejemplo ilustra la importancia de contar con este tipo de operaciones, y cómo el conocimiento de los agentes se ve afectado.

Ejemplo 8.1. Consideremos un escenario simplificado donde diferentes dos agentes, denotados como i y j respectivamente, desean hacer una torta. Ambos están intentando hacer una "buena" torta (representada por el símbolo proposicional g). Supongamos que ambos agentes están siguiendo una receta similar y que tienen todos los ingredientes a su disposición (h). La receta dice que para conseguir g se deben ejecutar los siguientes pasos o acciones: agregar huevos (e), batirlos (b), agregar harina (f), agregar leche (m), revolver estos ingredientes (s) y finalmente, cocinar la mezcla (p). De esta manera, el plan necesario para lograr g es $ebfmsp$. El agente i, que es un chef experimentado, es consciente de que esta es la forma de hacer una buena torta. Por otro lado, el agente j no tiene experiencia cocinando así que considera que el orden de ciertas instrucciones no es relevante. Supongamos en este caso que sea el orden de las acciones f y m ("agregar harina" y "agregar leche"). Es decir que, para este agente el plan correcto ($ebfmsp$) y el plan con el orden cambiado (*ebmfsp*) dan los mismos resultados. Dicha situación puede ser representada por el LTS^U $\mathfrak{M}$ que introducimos a continuación:

$$\mathfrak{M}: \begin{tikzcd} \boxed{h} \arrow[r, "e"] & \bigcirc \arrow[loop above, "b"] \arrow[r, "f"] & \bigcirc \arrow[r, "m"] & \bigcirc \arrow[loop above, "s"] \arrow[r, "p"] & \boxed{g} \end{tikzcd} \qquad \begin{array}{l} \mathrm{U}(i) = \{\{\mathit{ebfmsp}\}\} \\ \mathrm{U}(j) = \{\{\mathit{ebfmsp}, \mathit{ebmfsp}\}\} \end{array}$$

Dado que el agente i es un cocinero experimentado, $\mathrm{U}(i)$ tiene un único conjunto de planes $\{\mathit{ebfmsp}\}$, correspondiente a la receta correcta. Por lo tanto, $\mathfrak{M} \models \mathsf{Kh}_i(h, g)$. Mientras tanto, $\mathrm{U}(j)$ tiene el conjunto conformado por la receta original (*ebfmsp*) y la receta con las dos acciones intercambiadas (*ebmfsp*), que no es fuertemente ejecutable. Por lo tanto, $\mathfrak{M} \not\models \mathsf{Kh}_j(h, g)$. $\square$

Con el objeto de modelar la evolución o los cambios en el conocimiento de los agentes, tal como la situación ilustrada en el ejemplo anterior, se pueden definir modalidades dinámicas capaces de modificar tanto la información óntica como la información epistémica. Siguiendo esta idea, en las siguientes secciones se presentarán dos tipos de operadores dinámicos: ***(1)*** operadores ónticos y ***(2)*** operadores epistémicos. Los primeros están basados en operadores de anuncios públicos (PAL [17, 18, 91]) y de actualización de relaciones (AUL [69, 110]) de la lógica dinámica epistémica de 'saber que'. Por otro lado, los segundos refinan la relación de indistinguibilidad permitiendo distinguir ciertos planes de otros. También nos dedicaremos a discutir las propiedades, las consecuencias y desafíos a considerar con estos operadores, siendo la mayoría de estas desarrolladas en [9]. Por último, dado que analizaremos la expresividad de las siguientes extensiones, será necesario recordar la notación $\preceq$ introducida en la Definición 1.6, en este caso sobre modelos LTS^{U}s y generalizando para clases de modelos arbitrarias.

Definición 8.1 ($\mathsf{L}_1 \preceq_{\mathbf{M}} \mathsf{L}_2$)**.** Sean L_1 y L_2 dos lógicas, y $\mathbf{M}$ una clase de modelos LTS^{U}s. L_1 es tan expresiva como L_2 o reducible a L_2 en $\mathbf{M}$ (denotado como $\mathsf{L}_1 \preceq_{\mathbf{M}} \mathsf{L}_2$) si para toda fórmula φ de L_1 existe una fórmula φ' de L_2 tal que para todo modelo $\mathfrak{M} \in \mathbf{M}$ y $w \in \mathrm{D}_{\mathfrak{M}}$, se cumple que

$$\mathfrak{M}, w \models \varphi \;\text{ sii }\; \mathfrak{M}, w \models \varphi'.$$

Caso contrario, $\mathsf{L}_1 \not\preceq_{\mathbf{M}} \mathsf{L}_2$. Si $\mathsf{L}_1 \preceq_{\mathbf{M}} \mathsf{L}_2$ pero $\mathsf{L}_2 \not\preceq_{\mathbf{M}} \mathsf{L}_1$, se dice que L_2 es más expresiva que L_1 en $\mathbf{M}$ y se escribe $\mathsf{L}_1 \prec_{\mathbf{M}} \mathsf{L}_2$. Si $\mathsf{L}_1 \preceq_{\mathbf{M}} \mathsf{L}_2$ y $\mathsf{L}_2 \preceq_{\mathbf{M}} \mathsf{L}_1$, entonces son equivalentes en $\mathbf{M}$ y se lo denota como $\mathsf{L}_1 \equiv_{\mathbf{M}} \mathsf{L}_2$. Si la clase $\mathbf{M}$ es la de todos los LTS^{U}s, se la omite de la notación.

8.1. Operadores ónticos

8.1.1. Operador óntico basado en anuncios públicos

En esta subsección se introducirá un operador óntico basado en *anuncios públicos* (también llamadas observaciones públicas) para lógicas epistémicas de saber qué (PAL [91]). En la bibliografía, dicha modalidad representa una acción *epistémica* en la cual se anuncia a todos agentes una determinada información. Esto en la semántica se traduce como una modificación del modelo mediante la eliminación de estados (e indirectamente relaciones) que ya no cumplan con esta información. De esta manera, el agente ya no considera *posible* cierta información, y disminuye la incertidumbre sobre tales hechos. Sintácticamente la modalidad generalmente se denota como $[\chi]$. Considerando [109], la misma se interpreta de la siguiente manera.

Definición 8.2. Se extiende la relación $\models$ de la Definición 4.5 para el caso $[\chi]\varphi$:

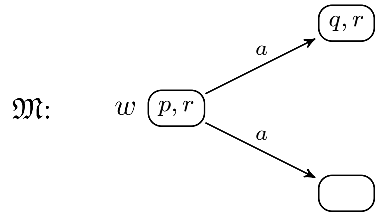
$$\mathfrak{M}, w \models [\chi]\varphi \;\text{ sii }\; \mathfrak{M}, w \models \chi, \text{ entonces } \mathfrak{M}_\chi, w \models \varphi,$$

con $\mathfrak{M}_\chi = \langle \mathrm{S}_\chi, \mathrm{R}_\chi, \mathrm{U}_\chi, \mathrm{V}_\chi \rangle$ donde

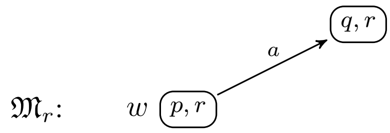
$$\begin{array}{ll} \mathrm{S}_\chi = \{w \mid \mathfrak{M}, w \models \chi\} = [\![\chi]\!]^{\mathfrak{M}} & (\mathrm{R}_\chi)_a = \mathrm{R}_a \cap \mathrm{S}_\chi^2, \text{ para cada } a \in \mathsf{Act} \\ \mathrm{V}_\chi(w) = \mathrm{V}(w), \text{ para todo } w \in \mathrm{S}_\chi & \mathrm{U}_\chi = \mathrm{U}. \end{array}$$

En esta definición, el operador elimina los mundos que no cumplen χ y restringe las relaciones y la función de valuación al nuevo dominio de $[\![\chi]\!]^{\mathfrak{M}}$. En la lógica de L_{Kh} de [116], las relaciones definen las habilidades del agente. Por lo tanto, esta acción de borrar mundos es un cambio tanto óntico como epistémico dado que las acciones disponibles para el agente cambian y con esto sus habilidades. Sin embargo, en la semántica basada en LTS^Us, las relaciones proveen información óntica únicamente. Notar que la definición de U_χ es provista aquí con el objeto de trabajar sobre la clase de modelos LTS^Us. Otras definiciones del mismo son posibles, pero no profundizaremos en ello en esta sección. Por lo tanto, sólo se produce un cambio óntico y ya no uno epistémico.

Ejemplo 8.2. Sea $\mathfrak{M}$ el siguiente LTS^U con un solo agente, con $\mathrm{U}(i) := \{\{a\}\}$:



Se puede ver que $\mathfrak{M}, w \models \neg\mathsf{Kh}_i(p, q)$ ya que el único plan disponible para el agente, a, alcanza un estado que no cumple q. Al anunciar r, el modelo resultante $\mathfrak{M}_r$ elimina dicho estado y se tiene que $\mathfrak{M}_r, w \models \mathsf{Kh}_i(p, q)$. Como además $\mathfrak{M}, w \models r$, por definición de $\models$, $\mathfrak{M}, w \models [r]\mathsf{Kh}_i(p, q)$.

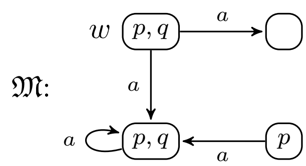


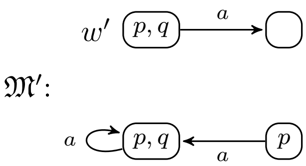
□

Utilizando un resultado similar al dado en [118] para la modalidad Kh con condiciones intermedias, se tiene que $[\chi]$ agrega expresividad a $\mathsf{L}_{\mathsf{Kh}_i}$.

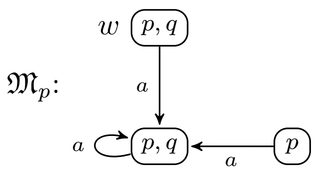
Proposición 8.1. *Sea* $\mathsf{L}_{\mathsf{Kh}_i} + [\chi]$ *la lógica* $\mathsf{L}_{\mathsf{Kh}_i}$ *extendida con* $[\chi]$*, se tiene que* $\mathsf{L}_{\mathsf{Kh}_i} \prec \mathsf{L}_{\mathsf{Kh}_i} + [\chi]$.

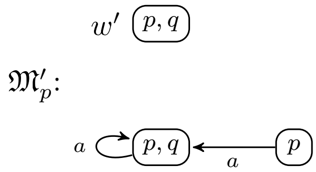
Demostración. Como $\mathsf{L}_{\mathsf{Kh}_i} + [\chi]$ es una extensión de $\mathsf{L}_{\mathsf{Kh}_i}$, se tiene que $\mathsf{L}_{\mathsf{Kh}_i} \preceq \mathsf{L}_{\mathsf{Kh}_i} + [\chi]$. Para probar que $\mathsf{L}_{\mathsf{Kh}_i} + [\chi] \not\preceq \mathsf{L}_{\mathsf{Kh}_i}$, se consideran los siguientes LTS^Us $\mathfrak{M}$ y $\mathfrak{M}'$, tales que $\mathrm{U}(i) = \mathrm{U}'(i) = \{\{a\}\}$, que son bisimilares en $\mathsf{L}_{\mathsf{Kh}_i}$ ($\mathfrak{M}, w \leftrightarroweq_{\mathsf{L}_{\mathsf{Kh}_i}} \mathfrak{M}', w'$). Usando el Teorema 6.1, $\mathfrak{M}, w \leftrightsquigarrow_{\mathsf{L}_{\mathsf{Kh}_i}} \mathfrak{M}', w'$ y por lo tanto cumplen las mismas fórmulas en $\mathsf{L}_{\mathsf{Kh}_i}$ (son indistinguibles en la lógica). En particular, ambos modelos no cumplen $\mathsf{Kh}_i(p, q)$.





Una vez anunciado p, todos los mundos que cumplen p se conservan, por ejemplo el estado w en $\mathfrak{M}$ y w' en $\mathfrak{M}'$. Más aún, $\mathsf{Kh}_i(p, q)$ se cumple en $\mathfrak{M}_p$ siendo $\{a\}$ el conjunto testigo. Esto se debe a que después el anuncio se descarta el mundo problemático. Con esto, $\mathfrak{M}, w \models p$ y $\mathfrak{M}_p, w \models \mathsf{Kh}_i(p, q)$. Por lo tanto, se cumple que $\mathfrak{M}, w \models [p]\mathsf{Kh}_i(p, q)$. Por otro lado, si bien $\mathfrak{M}', w' \models p$, se tiene que $\mathfrak{M}'_p, w \not\models \mathsf{Kh}_i(p, q)$ ya que el único conjunto de planes a considerar, $\{a\}$, no es fuertemente ejecutable en w', un estado que cumple p. Luego, por definición, $\mathfrak{M}', w' \not\models [p]\mathsf{Kh}_i(p, q)$.





Por lo tanto, existe una fórmula con la modalidad de anuncios públicos tal que distingue dos modelos bisimilares para $\mathsf{L}_{\mathsf{Kh}_i}$. Con esto, la lógica $\mathsf{L}_{\mathsf{Kh}_i}$ extendida con $[\chi]$ no es reducible a $\mathsf{L}_{\mathsf{Kh}_i}$ ($\mathsf{L}_{\mathsf{Kh}_i} + [\chi] \not\preceq \mathsf{L}_{\mathsf{Kh}_i}$) y, por ende, es más expresiva. □

El resultado anterior indica que la lógica $\mathsf{L}_{\mathsf{Kh}_i}$ no es lo suficientemente expresia para emular el comportamiento de la modalidad $[\chi]$, a diferencia del caso de la lógica epistémica básica (Definición 8.2). En tal caso, es posible eliminar la modalidad de anuncios mediante una codificación en la lógica básica [109]. Esto se debe a que la naturaleza del operador dinámico $[\chi]$ difiere bastante de la semántica del operador Kh_i. El mismo está diseñado para expresar implícitamente propiedades sobre la existencia de una manera de lograr un objetivo bajo ciertas circunstancias. A partir de esta comprensión sobre los efectos de un operador y de otro, podemos diseñar operadores dinámicos más acordes en el contexto de "saber cómo". A continuación presentaremos un operador como resultado de tales observaciones.

Definición 8.3. Las fórmulas del lenguaje $\mathsf{PAL}_{\mathsf{Kh}_i}$ están dadas por la siguiente gramática:

$$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid \mathsf{Kh}_i(\varphi, \varphi) \mid [!\varphi]\varphi,$$

con $p \in \mathsf{Prop}$ y $i \in \mathsf{Agt}$. Las constantes y demás conectores booleanos $\bot$, $\top$, $\wedge$, $\rightarrow$, y $\leftrightarrow$ se definen de la forma usual.

La siguiente definición alinea la semántica del operador Kh_i con la manera en que se actualiza el modelo mediante operadores dinámicos.

Definición 8.4. Sea $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \mathrm{U}, \mathrm{V} \rangle$ un LTS^U, χ una fórmula en $\mathsf{PAL}_{\mathsf{Kh}_i}$. Se define $\mathfrak{M}_{!\chi} = \langle \mathrm{S}_{!\chi}, \mathrm{R}_{!\chi}, \mathrm{U}_{!\chi}, \mathrm{V}_{!\chi} \rangle$, donde:

- $\mathrm{S}_{!\chi} = [\![\chi]\!]^{\mathfrak{M}}$,
- $(\mathrm{R}_{!\chi})_a = \{(w, v) \in \mathrm{R}_a \mid w \in [\![\chi]\!]^{\mathfrak{M}}, \mathrm{R}_a(w) \subseteq [\![\chi]\!]^{\mathfrak{M}}\}$ para cada $a \in \mathsf{Act}$,
- $\mathrm{U}_{!\chi} = \mathrm{U}$, y
- $\mathrm{V}_{!\chi}(w) = \mathrm{V}(w)$ (para cada $w \in \mathrm{S}_{!\chi}$).

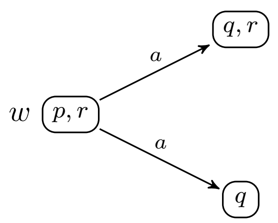
Se extiende la relación $\models$ de la Definición 4.5 para el caso $[!\chi]\varphi$:

$$\mathfrak{M}, w \models [!\chi]\varphi \quad \text{sii} \quad \mathfrak{M}, w \models \chi, \text{ entonces } \mathfrak{M}_{!\chi}, w \models \varphi.$$

La única diferencia entre $\mathfrak{M}_{!\chi}$ y la definición estándar de $\mathfrak{M}_{\chi}$, que es solamente la restricción de $\mathfrak{M}$ a los estados que satisfacen χ, es la definición de las relaciones. En este caso una condición más fuerte es necesaria para que una arista a de un estado $w \in [\![\chi]\!]^{\mathfrak{M}}$ persista después del cambio: si $\mathrm{R}_a(w) \not\subseteq [\![\chi]\!]^{\mathfrak{M}}$ entonces $(\mathrm{R}_{!\chi})_a(w) = \emptyset$, pero si $\mathrm{R}_a(w) \subseteq [\![\chi]\!]^{\mathfrak{M}}$ entonces $(\mathrm{R}_{!\chi})_a(w) = \mathrm{R}_a(w)$. Notar que en este contexto la eliminación de algunos estados indican que las situaciones descritas ya no son *alcanzables*, en vez de *posibles*, tal como lo expresado por Kh_i.

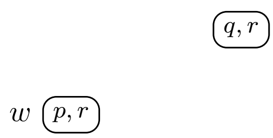
Ejemplo 8.3. Sea $\mathfrak{M}$ el siguiente LTS^U de un solo agente i, con $\mathrm{U}(i) := \{\{a\}\}$ del Ejemplo 8.2:

$\mathfrak{M}$:



Podemos observar que $\mathfrak{M}, w \models \mathsf{Kh}_i(p, q)$. Sin embargo, al anunciar r, el modelo resultante $\mathfrak{M}_{!r}$ elimina todos los estados que alcanza w vía a, ya que uno de estos no cumple r. Con esto, $\mathfrak{M}_{!r}, w \models \neg\mathsf{Kh}_i(p, q)$. Dado que $\mathfrak{M}, w \models r$, por definición de $\models$, $\mathfrak{M}, w \models [!r]\neg\mathsf{Kh}_i(p, q)$.

$\mathfrak{M}_{!r}$:



$\square$

Vale la pena destacar que estas dos formas de modificar el modelo se relacionan con las dos formas de actualzar modelos de "vecindad" (*neighbourhood models*) discutidas en [79]. Un modelo de vecindad [82, 97] está dado por un dominio S no vacío, una valuación atómica, y una función de vecindad $\mathrm{N} : \mathrm{S} \to 2^{2^{\mathrm{S}}}$, el cual asigna a un conjunto de estados a cada estado posible. Sea $U \subseteq \mathrm{S}$ un conjunto no vacío de estados. Por un lado, el submodelo actualizado vía *U-intersección*, definido en [79], tiene U como su dominio, con su función de vecindad construida mediante la restricción de cada conjunto en una vecindad al nuevo dominio. Este submodelo es análogo a lo que es $\mathfrak{M}_{\chi}$ con un anuncio público estándar. Por otro lado, el submodelo actualizado vía *U-subconjunto* [79] también tiene U como su dominio, pero su función de vecindad está construida a partir de mantener sólo aquellos conjuntos que ya son un subconjunto del nuevo dominio. Este submodelo es análogo a lo que es $\mathfrak{M}_{!\chi}$ con $[!\chi]$. Además de esto, $[!\chi]$ es una modalidad normal.

Proposición 8.2. *Teniendo en cuenta la definición de validez (Definición 4.5), sean φ_1, φ_2 y χ fórmulas de $\mathsf{PAL}_{\mathsf{Kh}_i}$, las siguientes propiedades se cumplen:*

1. $\models [!\chi](\varphi_1 \to \varphi_2) \to ([!\chi]\varphi_1 \to [!\chi]\varphi_2)$.
2. *Si* $\models \varphi$, *entonces* $\models [!\chi]\varphi$.

Demostración.

1. Sean $\mathfrak{M}$ un LTS^U y $w \in \mathrm{D}_{\mathfrak{M}}$ tales que $\mathfrak{M}, w \models [!\chi](\varphi_1 \to \varphi_2)$ y $\mathfrak{M} \models [!\chi]\varphi_1$. Supongamos que $\mathfrak{M}, w \models \chi$. Luego, por la definición de $\models$, se tiene que $\mathfrak{M}_{!\chi}, w \models (\varphi_1 \to \varphi_2)$ y $\mathfrak{M}_{!\chi}, w \models \varphi_1$. Nuevamente, por la definición de $\models$, $\mathfrak{M}_{!\chi}, w \models \varphi_2$ y con esto $\mathfrak{M}, w \models [!\chi]\varphi_2$. Como se tomaron $\mathfrak{M}$ y $w \in \mathrm{D}_{\mathfrak{M}}$ arbitrarios, $\models [!\chi](\varphi_1 \to \varphi_2) \to ([!\chi]\varphi_1 \to [!\chi]\varphi_2)$.

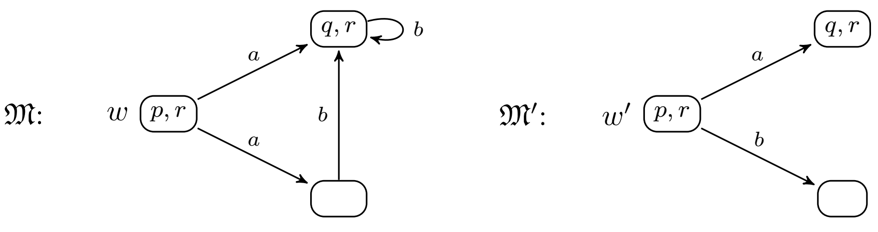
2. Supongamos que $\models \varphi$. Por definición, para todo LTS^U $\mathfrak{M}$ y estado $w \in \mathrm{D}_{\mathfrak{M}}$, se tiene que $\mathfrak{M}, w \models \varphi$. Sean $\mathfrak{M}$ un LTS^U y $w \in \mathrm{D}_{\mathfrak{M}}$ tales que $\mathfrak{M}, w \models \chi$. Dado que $\mathfrak{M}_{!\chi}$ es un LTS^U y $w \in \mathrm{D}_{\mathfrak{M}_{!\chi}}$, necesariamente $\mathfrak{M}_{!\chi}, w \models \varphi$ se cumple. Por lo tanto, por la Definición 8.4, $\mathfrak{M}, w \models [!\chi]\varphi$. Como se tomaron $\mathfrak{M}$ y $w \in \mathrm{D}_{\mathfrak{M}}$ arbitrarios, $\models [!\chi]\varphi$

□

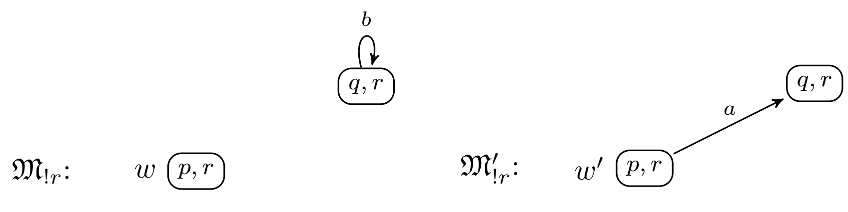
Sin embargo, aún con esta restricción en la definición, la lógica resultante sigue siendo más expresiva.

Proposición 8.3. $\mathsf{L}_{\mathsf{Kh}_i} \prec \mathsf{PAL}_{\mathsf{Kh}_i}$.

Demostración. Como $\mathsf{PAL}_{\mathsf{Kh}_i}$ es una extensión de $\mathsf{L}_{\mathsf{Kh}_i}$, se tiene que $\mathsf{L}_{\mathsf{Kh}_i} \preceq \mathsf{PAL}_{\mathsf{Kh}_i}$. Para probar que $\mathsf{PAL}_{\mathsf{Kh}_i} \not\preceq \mathsf{L}_{\mathsf{Kh}_i}$, se consideran los siguientes modelos de un solo agente $\mathfrak{M}$ y $\mathfrak{M}'$, con $\mathrm{U}(i) := \{\{ab\}\}$ y $\mathrm{U}'(i) := \{\{a\}\}$:



Los dos modelos son bisimilares en $\mathsf{L}_{\mathsf{Kh}_i}$ ($\mathfrak{M}, w \leftrightarroweq_{\mathsf{L}_{\mathsf{Kh}_i}} \mathfrak{M}', w'$) y por lo tanto satisfacen las mismas fórmulas en $\mathsf{L}_{\mathsf{Kh}_i}$ ($\mathfrak{M}, w \leftrightsquigarrow_{\mathsf{L}_{\mathsf{Kh}_i}} \mathfrak{M}', w'$). Sin embargo, $\mathfrak{M}, w \not\models [!r]\mathsf{Kh}_i(p, q)$ dado que $\mathfrak{M}, w \models r$ y $\mathfrak{M}_{!r}, w \not\models \mathsf{Kh}_i(p, q)$, mientras que $\mathfrak{M}', w' \models [!r]\mathsf{Kh}_i(p, q)$ dado que $\mathfrak{M}', w' \models r$ y $\mathfrak{M}'_{!r}, w \models \mathsf{Kh}_i(p, q)$.



Con esto, $\mathsf{PAL}_{\mathsf{Kh}_i} \not\preceq \mathsf{L}_{\mathsf{Kh}_i}$ y por lo tanto, $\mathsf{L}_{\mathsf{Kh}_i} \prec \mathsf{PAL}_{\mathsf{Kh}_i}$. □

RAtom	$\vdash [!\chi]p \leftrightarrow (\chi \to p)$
R$\neg$	$\vdash [!\chi]\neg\varphi \leftrightarrow (\chi \to \neg[!\chi]\varphi)$
R$\vee$	$\vdash [!\chi](\varphi \vee \psi) \leftrightarrow [!\chi]\varphi \vee [!\chi]\psi$
RKh	$\vdash [!\chi]\mathsf{Kh}_i(\psi, \varphi) \leftrightarrow (\chi \to \mathsf{Kh}_i(\chi \wedge [!\chi]\psi, \chi \wedge [!\chi]\varphi))$
RE$_{[!]}$	De $\vdash \varphi \leftrightarrow \psi$ se deriva $\vdash [!\chi]\varphi \leftrightarrow [!\chi]\psi$

Tabla 8.1: Axiomas de reducción $\mathcal{L}_{\mathsf{PAL}_{\mathsf{Kh}_i}}$.

Como mencionamos, algunos operadores dinámicos, tal como los anuncios públicos en lógica epistémica clásica (Ejemplo 2.3), no agregan expresividad y por lo tanto es posible axiomatizarlos mediante axiomas de reducción (ver Teorema 2.2). Los axiomas de reducción son interesantes ya que usualmente nos proveen una manera clara de comprender (de manera estática) el comportamiento de un operador. En este sentido, el resultado anterior no parece alentador. Sin embargo, existen otras posibilidades como por ejemplo considerar clases de modelos más simples que aún sean de interés. Veremos que sobre algunas de estas clases, es posible definir axiomas de reducción.

Dado un LTS^Us, para cada agente i, el conjunto $\mathrm{U}(i)$ determina la percepción del agente con respecto a sus habilidades. Por ejemplo, puede suceder que dos planes, ab y cd, sean indistinguibles para un agente i (es decir que pertenezcan a un mismo $\pi \in \mathrm{U}(i)$). Específicamente, del final del Capítulo 5 se puede concluir que la lógica no puede distinguir entre la clase arbitraria de LTS^Us y la clase de LTS^Us donde cada $\pi \in \mathrm{U}(i)$ es un conjunto unitario tal que $\pi \subseteq \mathsf{Act}$. Esto ya no se cumple con la introducción de $[!\chi]$, dado que la eliminación de los estados potencialmente descarta testigos para las modalidades Kh_i, como sucedía con los modelos de la Proposición 8.3. Teniendo en cuenta esto, consideremos una clase particular de modelos.

Definición 8.5. Se define la clase de modelos LTS^Us tales que sus agentes consideran únicamente acciones básicas (**BA**):

$$\mathbf{M_{BA}} := \{\mathfrak{M} \mid \mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \mathrm{U}, \mathrm{V}\rangle \text{ es un } \mathrm{LTS}^U \text{ y para cada } i \in \mathsf{Agt}, \text{ si } \pi \in \mathrm{U}(i), \text{ entonces } \pi \subseteq \mathsf{Act}\}.$$

La clase $\mathbf{M_{BA}}$ puede corresponderse, por ejemplo, a una representación más abstracta de las habilidades de los agentes, para los cuales un curso de acción está modelado como una acción simple. Este tipo de representación simplificada es usual en el modelado de sistemas multiagente y planning (ver por ejemplo la idea de abstracción en planning [49]). Como se verá a continuación, los axiomas de reducción de la Tabla 8.1 son válidos sobre la clase de modelos $\mathbf{M_{BA}}$ teniendo en cuenta un lema auxiliar (Lema 8.1).

Lema 8.1. *Sean φ y χ fórmulas de $\mathsf{PAL}_{\mathsf{Kh}_i}$, las siguientes igualdades se cumplen:*

1. $[\![[!\chi]\varphi]\!]^{\mathfrak{M}} = [\![\neg\chi]\!]^{\mathfrak{M}} \cup [\![\varphi]\!]^{\mathfrak{M}_{!\chi}}$ *y*
2. $[\![\varphi]\!]^{\mathfrak{M}_{!\chi}} = [\![\chi \wedge [!\chi]\varphi]\!]^{\mathfrak{M}}$.

Demostración.

1. Sea $w \in [\![[!\chi]\varphi]\!]^{\mathfrak{M}}$, por definición esto equivale a que $\mathfrak{M}, w \models [!\chi]\varphi$. Por ende, $\mathfrak{M}, w \models \chi$ implica $\mathfrak{M}_{!\chi}, w \models \varphi$. Dicha implicación puede ser reescrita como $\mathfrak{M}, w \models \neg\chi$ ó $\mathfrak{M}_{!\chi}, w \models \varphi$. Por lo tanto, $w \in [\![\neg\chi]\!]^{\mathfrak{M}}$ ó $[\![\varphi]\!]^{\mathfrak{M}_{!\chi}}$. Lo cual equivale a que $w \in [\![\neg\chi]\!]^{\mathfrak{M}} \cup [\![\varphi]\!]^{\mathfrak{M}_{!\chi}}$.

2. Sea $w \in [\![\varphi]\!]^{\mathfrak{M}_{!\chi}}$, por definición $w \in \mathrm{S}_{!\chi} = [\![\chi]\!]^{\mathfrak{M}}$ y $w \in [\![\varphi]\!]^{\mathfrak{M}_{!\chi}}$. Esto equivale a que $w \in [\![\chi]\!]^{\mathfrak{M}} \cap [\![\varphi]\!]^{\mathfrak{M}_{!\chi}}$. El conjunto anterior es de la forma $(A \cap B)$. Utilizando un razonamiento de conjuntos, esto es lo mismo que $(A \cap (A^c \cup B))$. Es decir, es equivalente a $w \in [\![\chi]\!]^{\mathfrak{M}} \cap ([\![\neg\chi]\!]^{\mathfrak{M}} \cup [\![\varphi]\!]^{\mathfrak{M}_{!\chi}})$. Tomando el resultado del primer ítem, $w \in [\![\chi]\!]^{\mathfrak{M}} \cap [\![[!\chi]\varphi]\!]^{\mathfrak{M}}$ y con esto, $w \in [\![\chi \wedge [!\chi]\varphi]\!]^{\mathfrak{M}}$.

□

Gracias a este lema, podemos obtener el siguiente resultado.

Lema 8.2. *Los axiomas de reducción de la Tabla 8.1 son válidos en* $\mathbf{M_{BA}}$.

Demostración. La regla $\mathsf{RE}_{[!]}$ se cumple aplicando la Proposición 8.2 y MP de $\mathcal{L}^{\mathrm{LTS}^U}_{\mathsf{Kh}_i}$.

- Para RAtom, si $\mathfrak{M}, w \models \chi$, entonces $\mathfrak{M}_{!\chi}, w \models p$ sii $p \in \mathrm{V}_{!\chi}(w)$. Como $\mathrm{V}_{!\chi}(w) = \mathrm{V}(w)$, se tiene que $\mathfrak{M}, w \models p$. Por definición, $\mathfrak{M}, w \models [!\chi]p$ sii $\mathfrak{M}, w \models \chi$ implica $\mathfrak{M}_{!\chi}, w \models p$, que es equivalente a que si $\mathfrak{M}, w \models \chi$, entonces $\mathfrak{M}, w \models p$. Por lo tanto es equivalente a $\mathfrak{M}, w \models (\chi \to p)$.

- Para $\mathsf{R}\neg$, por definición $\mathfrak{M}, w \models (\chi \to \neg[!\chi]\varphi)$ sii $\mathfrak{M}, w \models \chi$ implica $\mathfrak{M}, w \models \neg[!\chi]\varphi$. Usando de vuelta definición de $\models$, $\mathfrak{M}, w \models \neg[!\chi]\varphi$ sii $\mathfrak{M}, w \models \chi$ y $\mathfrak{M}_{!\chi}, w \not\models \varphi$. Por lo tanto, $\mathfrak{M}, w \models (\chi \to \neg[!\chi]\varphi)$ sii $\mathfrak{M}, w \models \chi$ implica $\mathfrak{M}, w \models \chi$ y $\mathfrak{M}_{!\chi}, w \not\models \varphi$. La anterior implicación es de la forma $(A \to (A \wedge B))$. La cual es equivalente a $(A \to B)$. Es decir que $\mathfrak{M}, w \models \chi$ implica $\mathfrak{M}_{!\chi}, w \not\models \varphi$. Como $\mathfrak{M}_{!\chi}, w \not\models \varphi$ sii $\mathfrak{M}_{!\chi}, w \models \neg\varphi$, por la definición de $[!\chi]$, se tiene que $\mathfrak{M}, w \models [!\chi]\neg\varphi$.

- Para $\mathsf{R}\vee$, por definición de $[!\chi]$ se tiene que $\mathfrak{M}, w \models [!\chi](\varphi \vee \psi)$ sii $\mathfrak{M}, w \models \chi$ implica $\mathfrak{M}_{!\chi}, w \models \varphi \vee \psi$. Esto es equivalente a $\mathfrak{M}, w \models \chi$ implica $\mathfrak{M}_{!\chi}, w \models \varphi$ ó $\mathfrak{M}_{!\chi}, w \models \psi$. La anterior implicación es de la forma $(A \to (B \vee C))$. Lo cual equivale a $((A \to B) \vee (A \to C))$. Es decir, $\mathfrak{M}, w \models \chi$ implica $\mathfrak{M}_{!\chi}, w \models \varphi$, ó $\mathfrak{M}, w \models \chi$ implica $\mathfrak{M}_{!\chi}, w \models \psi$. Usando la definición de $[!\chi]$, $\mathfrak{M}, w \models [!\chi]\varphi$ ó $\mathfrak{M}, w \models [!\chi]\psi$ y por lo tanto $\mathfrak{M}, w \models [!\chi]\varphi \vee [!\chi]\psi$.

- Para RKh, por la definición de $[!\chi]$, $\mathfrak{M}, w \models [!\chi]\mathsf{Kh}_i(\psi, \varphi)$ sii $\mathfrak{M}, w \models \chi$ implica $\mathfrak{M}_{!\chi}, w \models \mathsf{Kh}_i(\psi, \varphi)$. Usando la definición de Kh_i, $\mathfrak{M}_{!\chi}, w \models \mathsf{Kh}_i(\psi, \varphi)$ sii existe $\pi \in (\mathrm{U}_{!\chi})(i)$ con $\pi \subseteq \mathsf{Act}$ tal que $[\![\psi]\!]^{\mathfrak{M}_{!\chi}} \subseteq \mathrm{SE}^{\mathfrak{M}_{!\chi}}(\pi)$ y $(\mathrm{R}_{!\chi})_\pi([\![\psi]\!]^{\mathfrak{M}_{!\chi}}) \subseteq [\![\varphi]\!]^{\mathfrak{M}_{!\chi}}$. Por la definición de $\mathfrak{M}_{!\chi}$ y el Lema 8.1, $(\mathrm{U}_{!\chi})(i) = \mathrm{U}(i)$ y $[\![\psi]\!]^{\mathfrak{M}_{!\chi}} = [\![\chi \wedge [!\chi]\psi]\!]^{\mathfrak{M}}$ y $[\![\varphi]\!]^{\mathfrak{M}_{!\chi}} = [\![\chi \wedge [!\chi]\varphi]\!]^{\mathfrak{M}}$. Por lo tanto, $\mathfrak{M}_{!\chi}, w \models \mathsf{Kh}_i(\psi, \varphi)$ sii existe $\pi \in \mathrm{U}(i)$ con $\pi \subseteq \mathsf{Act}$ tal que $[\![\chi \wedge [!\chi]\psi]\!]^{\mathfrak{M}} \subseteq \mathrm{SE}^{\mathfrak{M}_{!\chi}}(\pi)$ y $(\mathrm{R}_{!\chi})_\pi([\![\chi \wedge [!\chi]\psi]\!]^{\mathfrak{M}}) \subseteq [\![\chi \wedge [!\chi]\varphi]\!]^{\mathfrak{M}}$. Sea $a \in \pi$ y $w \in \mathrm{S}_{!\chi}$. Si $w \in [\![\chi \wedge [!\chi]\psi]\!]^{\mathfrak{M}}$, entonces:

  - $(\mathrm{R}_{!\chi})_a(w) \neq \emptyset$ (dado que $w \in \mathrm{SE}^{\mathfrak{M}_{!\chi}}(a)$) y
  - $(\mathrm{R}_{!\chi})_a(w) \subseteq [\![\chi \wedge [!\chi]\varphi]\!]^{\mathfrak{M}}$.

  Usando la Definición 8.4, el primer ítem es equivalente a $w \in [\![\chi]\!]^{\mathfrak{M}}$, $\mathrm{R}_a(w) \subseteq [\![\chi]\!]^{\mathfrak{M}}$ y $\mathrm{R}_a(w) \neq \emptyset$. Con esto, $(\mathrm{R}_{!\chi})_a(w) = \mathrm{R}_a(w)$ que se aplica al segundo ítem. Por lo tanto, si $w \in [\![\chi \wedge [!\chi]\psi]\!]^{\mathfrak{M}}$, entonces:

  - $w \in [\![\chi]\!]^{\mathfrak{M}}$, $\mathrm{R}_a(w) \subseteq [\![\chi]\!]^{\mathfrak{M}}$, $\mathrm{R}_a(w) \neq \emptyset$ y
  - $\mathrm{R}_a(w) \subseteq [\![\chi \wedge [!\chi]\varphi]\!]^{\mathfrak{M}}$.

  Luego $w \in [\![\chi]\!]^{\mathfrak{M}}$ y $\mathrm{R}_a(w) \subseteq [\![\chi]\!]^{\mathfrak{M}}$ son redundantes dado que $w \in [\![\chi \wedge [!\chi]\psi]\!]^{\mathfrak{M}}$ y $\mathrm{R}_a(w) \subseteq [\![\chi \wedge [!\chi]\varphi]\!]^{\mathfrak{M}}$. Con esto, si $w \in [\![\chi \wedge [!\chi]\psi]\!]^{\mathfrak{M}}$, entonces:

  - $\mathrm{R}_a(w) \neq \emptyset$ (por lo tanto, $w \in \mathrm{SE}^{\mathfrak{M}}(a)$) y
  - $\mathrm{R}_a(w) \subseteq [\![\chi \wedge [!\chi]\varphi]\!]^{\mathfrak{M}}$.

  Como se demostró para $a \in \pi$ y $w \in [\![\chi \wedge [!\chi]\psi]\!]^{\mathfrak{M}}$ arbitrarios, el resultado se cumple para todo $a \in \pi$ y $w \in [\![\chi \wedge [!\chi]\psi]\!]^{\mathfrak{M}}$. Luego, $\mathfrak{M}_{!\chi}, w \models \mathsf{Kh}_i(\psi, \varphi)$ sii existe $\pi \in \mathrm{U}(i)$ con $\pi \subseteq \mathsf{Act}$ tal que $[\![\chi \wedge [!\chi]\psi]\!]^{\mathfrak{M}} \subseteq \mathrm{SE}(\pi)$ y $\mathrm{R}_\pi([\![\chi \wedge [!\chi]\psi]\!]^{\mathfrak{M}}) \subseteq [\![\chi \wedge [!\chi]\varphi]\!]^{\mathfrak{M}}$. Esto sucede sii $\mathfrak{M}, w \models \mathsf{Kh}_i(\chi \wedge [!\chi]\psi, \chi \wedge [!\chi]\varphi)$. Como esta equivalencia se prueba asumiendo $\mathfrak{M}, w \models \chi$, entonces $\mathfrak{M}, w \models \chi$ implica $\mathfrak{M}_{!\chi}, w \models \mathsf{Kh}_i(\psi, \varphi)$ sii $\mathfrak{M}, w \models \chi$ implica $\mathfrak{M}, w \models \mathsf{Kh}_i(\chi \wedge [!\chi]\psi, \chi \wedge [!\chi]\varphi)$. Por lo tanto, es equivalente a $\mathfrak{M}, w \models \chi \to \mathsf{Kh}_i(\chi \wedge [!\chi]\psi, \chi \wedge [!\chi]\varphi)$.

□

Con este resultado, dada una fórmula de $\mathsf{PAL}_{\mathsf{Kh}_i}$, se pueden eliminar las modalidades $[!\chi]$, obteniendo así una fórmula equivalente de $\mathsf{L}_{\mathsf{Kh}_i}$. Por lo tanto, $\mathsf{PAL}_{\mathsf{Kh}_i}$ y $\mathsf{L}_{\mathsf{Kh}_i}$ son equivalentes en los modelos de $\mathbf{M_{BA}}$.

Corolario 8.1. $\mathsf{L}_{\mathsf{Kh}_i} \equiv_{\mathbf{M_{BA}}} \mathsf{PAL}_{\mathsf{Kh}_i}$.

Con esto, se tiene que el sistema axiomático $\mathcal{L}^{\mathrm{LTS}^U}_{\mathsf{Kh}_i}$ más los axiomas de reducción son correctos y fuertemente completos con respecto a la clase $\mathbf{M_{BA}}$.

Teorema 8.1. *$\mathcal{L}^{\mathrm{LTS}^U}_{\mathsf{Kh}_i}$ más los axiomas de reducción para $[!\chi]$ en la Tabla 8.1 son correctos y fuertemente completos para $\mathsf{PAL}_{\mathsf{Kh}_i}$ con respecto a la clase de modelos $\mathbf{M_{BA}}$.*

Demostración. Correctitud es directa del Teorema 5.1 y el Lema 8.2. Como $\mathsf{PAL}_{\mathsf{Kh}_i}$ es reducible a $\mathsf{L}_{\mathsf{Kh}_i}$ en $\mathbf{M_{BA}}$, sea Ω un conjunto consistente en $\mathsf{PAL}_{\mathsf{Kh}_i}$, el mismo es equivalente a otro conjunto consistente Γ' en $\mathsf{L}_{\mathsf{Kh}_i}$ para $\mathbf{M_{BA}}$. Usando la prueba del Teorema 5.1, siendo Γ la extensión maximal consistente de Γ', Γ' es satisfacible en el modelo canónico $\mathfrak{M}^{\Gamma}$. Por la Definición 5.3, $\mathfrak{M}^{\Gamma}$ es un modelo de la clase $\mathbf{M_{BA}}$ y por lo tanto Ω es satisfacible. Con esto se prueba que todo conjunto consistente Ω en $\mathsf{PAL}_{\mathsf{Kh}_i}$ es satisfacible por un modelo de la clase $\mathbf{M_{BA}}$. □

Para terminar esta sección, se tiene que el problema de satisfacibilidad para $\mathsf{PAL}_{\mathsf{Kh}_i}$ es decidible al menos en $\mathbf{M_{BA}}$. Esto quiere decir que, dada una fórmula φ de $\mathsf{PAL}_{\mathsf{Kh}_i}$, existe un procedimiento que en una cantidad finita de pasos determina si la misma es satisfacible o no en $\mathbf{M_{BA}}$.

Corolario 8.2. *El problema de satisfacibilidad para* $\mathsf{PAL}_{\mathsf{Kh}_i}$ *sobre* $\mathbf{M_{BA}}$ *es decidible.*

Demostración. Dada una fórmula φ de $\mathsf{PAL}_{\mathsf{Kh}_i}$ arbitraria, la misma se la puede traducir en una cantidad finita de pasos a una fórmula equivalente φ' de $\mathsf{L}_{\mathsf{Kh}_i}$ en $\mathbf{M_{BA}}$ (Corolario 8.1). Como el problema de satisfacibilidad para $\mathsf{L}_{\mathsf{Kh}_i}$ es decidible (Teorema 7.2), se tiene un procedimiento, si bien no determinístico, que en una cantidad finita de pasos determina si φ' es satisfacible o no. Luego de dicho procedimiento, se tienen dos resultados posibles: ***(1)*** Si φ' es satisfacible, entonces lo es en $\mathbf{M^F_{NU}}$ (Corolario 7.2). Como $\mathbf{M^F_{NU}} \subseteq \mathbf{M_{BA}}$, dado que cada $\mathrm{U}(i)$ consta de conjuntos unitarios de acciones, se tiene que φ es satisfacible en $\mathbf{M_{BA}}$. ***(2)*** Si φ' no es satisfacible, entonces tampoco lo es en $\mathbf{M_{BA}}$. Por el Corolario 8.1, φ no es satisfacible en $\mathbf{M_{BA}}$. Por lo tanto, se tiene un procedimiento tal que resuelve el problema de satisfacibilidad para $\mathsf{PAL}_{\mathsf{Kh}_i}$ en $\mathbf{M_{BA}}$. □

8.1.2. Operador óntico basado en actualización de relaciones

Así como en PAL usualmente se eliminan estados del modelo, otra forma de modificar la información óntica del mismo es conservando los estados pero eliminando las aristas. Este enfoque está basado en la lógica de actualización de relaciones (*Arrow Update Logic*, AUL; [69]). En lógica epistémica estándar de saber que, esto se corresponde con cambios en la incertidumbre del agente (intuitivamente, al reducir la indistinguibilidad entre mundos, el agente gana conocimiento). Para lógicas de saber cómo la situación es diferente: modificar aristas en un LTS corresponde a modificar las habilidades para los agentes, dado que las aristas representan la ejecución de acciones. Con esto en mente, se presenta la siguiente lógica con un operador de AUL en el contexto de $\mathsf{L}_{\mathsf{Kh}_i}$.

Definición 8.6. Las fórmulas del lenguaje $\mathsf{AUL}_{\mathsf{Kh}_i}$ están dadas por la siguiente gramática:

$$\begin{array}{lcl} \varphi & ::= & p \mid \neg\varphi \mid \varphi \vee \varphi \mid \mathsf{Kh}_i(\varphi,\varphi) \mid [E]\varphi, \\ E & ::= & (\varphi,\varphi) \mid E,(\varphi,\varphi), \end{array}$$

con $p \in \mathsf{Prop}$ y $i \in \mathsf{Agt}$.

Definición 8.7. Sea $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \mathrm{U}, \mathrm{V} \rangle$ un LTS^U, y $E = (\theta_1, \theta'_1), \ldots, (\theta_n, \theta'_n)$ tal que θ_i, θ'_i son fórmulas de $\mathsf{AUL}_{\mathsf{Kh}_i}$, para todo $0 \leq i \leq n$. Se define $\mathfrak{M}_E = \langle \mathrm{S}, \mathrm{R}_E, \mathrm{U}, \mathrm{V} \rangle$, donde para cada $a \in \mathsf{Act}$,

$$(\mathrm{R}_E)_a = \{(w,v) \in \mathrm{R}_a(w) \mid w \in [\![\bigwedge_{i=1}^{n} \theta_i]\!]^{\mathfrak{M}},\ \mathrm{R}_a(w) \subseteq [\![\bigwedge_{i=1}^{n} \theta'_i]\!]^{\mathfrak{M}}\}.$$

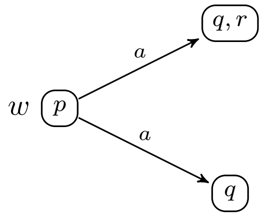
Notar que si $w \in [\![\bigwedge_{i=1}^{n} \theta_i]\!]^{\mathfrak{M}}$ y $\mathrm{R}_a(w) \subseteq [\![\bigwedge_{i=1}^{n} \theta'_i]\!]^{\mathfrak{M}}$, entonces $(\mathrm{R}_E)_a(w) = \mathrm{R}_a(w)$. Más aún, $(\mathrm{R}_E)_a(w) \neq \emptyset$ si y sólo si $w \in [\![\bigwedge_{i=1}^{n} \theta_i]\!]^{\mathfrak{M}}$, $\mathrm{R}_a(w) \subseteq [\![\bigwedge_{i=1}^{n} \theta'_i]\!]^{\mathfrak{M}}$ y $\mathrm{R}_a(w) \neq \emptyset$. Una vez más, la semántica de esta modalidad difiere de la original [69]. Las aristas para una determinada acción se conservan sólo si el estado de origen satisface la precondición y todos los estados que alcanza cumplen la postcondición.

Definición 8.8. Se extiende la relación de $\models$ de la Definición 4.5 con el caso para $[E]$:

$$\mathfrak{M}, w \models [E]\varphi \ \text{ sii } \ \mathfrak{M}_E, w \models \varphi.$$

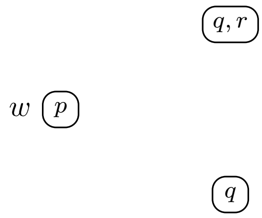
Ejemplo 8.4. Sea $\mathfrak{M}$ el siguiente LTS^U de un solo agente con $\mathrm{U}(i) := \{\{a\}\}$:

$\mathfrak{M}$:



Si bien, $\mathfrak{M}, w \models \mathsf{Kh}_i(p,q)$, al actualizar las aristas con $[(p,r)]$, en el modelo resultante $\mathfrak{M}_{(p,r)}$ se eliminan todas las aristas etiquetadas por a que comienzan desde w, ya que uno de estos no cumple r. Por otro lado, se conservan todos los estados. Con esto, $\mathfrak{M}_{(p,r)}, w \models \neg\mathsf{Kh}_i(p,q)$ y por definición de $\models$, $\mathfrak{M}, w \models [(p,r)]\neg\mathsf{Kh}_i(p,q)$.

$\mathfrak{M}_{(p,r)}$:



□

Como sucede con $\mathsf{PAL}_{\mathsf{Kh}_i}$, $\mathsf{AUL}_{\mathsf{Kh}_i}$ realiza sólo cambios en la información óntica a nivel de relaciones en la semántica basada en LTS^U. Como propiedades de la modalidad $[E]$, se tiene que el dual, $\langle E\rangle$, definido como $\langle E\rangle\varphi = \neg[E]\neg\varphi$ es equivalente a la modalidad $[E]$. Más aún, como en la Proposición 8.4, la modalidad $[E]$ es normal.

Proposición 8.4. *Sean $E = (\theta_1, \theta_1'), \ldots, (\theta_n, \theta_n')$, φ, φ_1, φ_2 fórmulas de $\mathsf{AUL}_{\mathsf{Kh}_i}$, las siguientes propiedades se cumplen:*

1. $\models [E]\varphi \leftrightarrow \neg[E]\neg\varphi$.
2. $\models [E](\varphi_1 \to \varphi_2) \to ([E]\varphi_1 \to [E]\varphi_2)$.
3. *Si* $\models \varphi$, *entonces* $\models [E]\varphi$.

Demostración. Sean $\mathfrak{M}$ un LTS^U y $w \in \mathrm{D}_{\mathfrak{M}}$, se dan las siguientes equivalencias: $\mathfrak{M}, w \models \neg[E]\neg\varphi$ sii $\mathfrak{M}, w \not\models [E]\neg\varphi$ sii $\mathfrak{M}_E, w \not\models \neg\varphi$ sii $\mathfrak{M}_E, w \models \varphi$ sii $\mathfrak{M}, w \models [E]\varphi$. Los demás ítems se prueban de manera similar a la Proposición 8.2. □

Por otro lado, $\mathsf{AUL}_{\mathsf{Kh}_i}$ no es reducible a $\mathsf{L}_{\mathsf{Kh}_i}$ para la clase de todos los modelos.

Proposición 8.5. $\mathsf{L}_{\mathsf{Kh}_i} \prec \mathsf{AUL}_{\mathsf{Kh}_i}$.

Demostración. Como $\mathsf{AUL}_{\mathsf{Kh}_i}$ es una extensión de $\mathsf{L}_{\mathsf{Kh}_i}$, se tiene que $\mathsf{L}_{\mathsf{Kh}_i} \preceq \mathsf{AUL}_{\mathsf{Kh}_i}$. Usando los modelos de la Proposición 8.3, se tiene que $\mathfrak{M}, w \not\models [(r,r)]\mathsf{Kh}_i(p,q)$ y $\mathfrak{M}', w' \models [(r,r)]\mathsf{Kh}_i(p,q)$. Con esto, $\mathsf{AUL}_{\mathsf{Kh}_i} \not\preceq \mathsf{L}_{\mathsf{Kh}_i}$ y por ende, $\mathsf{L}_{\mathsf{Kh}_i} \prec \mathsf{AUL}_{\mathsf{Kh}_i}$. □

Tal como ocurría con $\mathsf{PAL}_{\mathsf{Kh}_i}$, para esta lógica existen axiomas de reducción (Tabla 8.2) que son válidos para la clase de modelos $\mathbf{M_{BA}}$, y se pueden usar para eliminar todas las ocurrencias de la modalidad $[E]$. De manera análoga a la Tabla 8.1, se prueba que los axiomas de reducción de la Tabla 8.2 son válidos.

Lema 8.3. *Los axiomas de reducción de la Tabla 8.2 son válidos en* $\mathbf{M_{BA}}$.

Demostración. La regla RE_U se cumple aplicando la Proposición 8.4 y MP de $\mathcal{L}_{\mathsf{Kh}_i}^{\mathrm{LTS}^U}$.

- Para RJoin, sean $E' = (\bigwedge_{i=1}^n \theta_i, \bigwedge_{i=1}^n \theta_i') = (\theta, \theta')$ y $a \in \mathsf{Act}$, $(w,v) \in (\mathrm{R}_E)_a$ sii $w \in [\![\bigwedge_{i=1}^n \theta_i]\!]^{\mathfrak{M}}$, $\mathrm{R}_a(w) \subseteq [\![\bigwedge_{i=1}^n \theta_i']\!]^{\mathfrak{M}}$. Como $[\![\bigwedge_{i=1}^n \theta_i]\!]^{\mathfrak{M}} = [\![\theta]\!]^{\mathfrak{M}}$ y $[\![\bigwedge_{i=1}^n \theta_i']\!]^{\mathfrak{M}} = [\![\theta']\!]^{\mathfrak{M}}$, $(w,v) \in (\mathrm{R}_E)_a$ sii $(w,v) \in (\mathrm{R}_{E'})_a$. Con esto, $(\mathrm{R}_E)_a = (\mathrm{R}_{E'})_a$ para todo $a \in \mathsf{Act}$, $\mathfrak{M}_E = \mathfrak{M}_{E'}$ y satisfacen las mismas fórmulas. Por lo tanto, $\mathfrak{M}, w \models [E]\varphi$ sii $\mathfrak{M}, w \models [E']\varphi$.
- Para RAtom, sea $E = (\theta, \theta')$, $\mathfrak{M}, w \models [E]p$ sii $\mathfrak{M}_E, w \models p$. Esto equivale a $p \in \mathrm{V}(w)$ y con esto $\mathfrak{M}, w \models [E]p$ sii $\mathfrak{M}, w \models p$.

RJoin	$[E]\varphi \leftrightarrow [(\bigwedge_{i=1}^{n} \theta_i, \bigwedge_{i=1}^{n} \theta'_i)]\varphi$
RAtom	$[(\theta,\theta')]p \leftrightarrow p$
R$\neg$	$[(\theta,\theta')]\neg\varphi \leftrightarrow \neg[(\theta,\theta')]\varphi$
R$\vee$	$[(\theta,\theta')](\varphi \vee \psi) \leftrightarrow [(\theta,\theta')]\varphi \vee [(\theta,\theta')]\psi$
RKh	$[(\theta,\theta')]\mathsf{Kh}_i(\varphi,\psi) \leftrightarrow \mathsf{A}([(\theta,\theta')]\varphi \to \theta) \wedge \mathsf{Kh}_i([(\theta,\theta')]\varphi, \theta' \wedge [(\theta,\theta')]\psi)$
RE$_U$	De $\vdash \varphi \leftrightarrow \psi$ se deriva $\vdash [(\theta,\theta')]\varphi \leftrightarrow [(\theta,\theta')]\psi$

Tabla 8.2: Axiomas de reducción $\mathcal{L}_{\mathsf{AUL}_{\mathsf{Kh}_i}}$ con $E = (\theta_1, \theta'_1), \ldots, (\theta_n, \theta'_n)$.

- Para R$\neg$, sea $E = (\theta, \theta')$, $\mathfrak{M}, w \models [E]\neg\varphi$ sii $\mathfrak{M}_E, w \models \neg\varphi$ sii $\mathfrak{M}_E, w \not\models \varphi$ sii $\mathfrak{M}, w \not\models [E]\varphi$ sii $\mathfrak{M}, w \models \neg[E]\varphi$.

- Para R$\vee$, sea $E = (\theta, \theta')$, $\mathfrak{M}, w \models [E](\varphi \vee \psi)$ sii $\mathfrak{M}_E, w \models \varphi \vee \psi$ sii $\mathfrak{M}_E, w \models \varphi$ ó $\mathfrak{M}_E, w \models \psi$ sii $\mathfrak{M}, w \models [E]\varphi$ ó $\mathfrak{M}, w \models [E]\psi$ sii $\mathfrak{M}, w \models [E]\varphi \vee [E]\psi$.

- Para RKh, sea $E = (\theta, \theta')$, por la definición de $[E]$, $\mathfrak{M}, w \models [E]\mathsf{Kh}_i(\varphi, \psi)$ sii $\mathfrak{M}_E, w \models \mathsf{Kh}_i(\varphi, \psi)$. Usando la definición de Kh_i, $\mathfrak{M}_E, w \models \mathsf{Kh}_i(\varphi, \psi)$ sii existe $\pi \in \mathrm{U}(i)$ con $\pi \subseteq \mathsf{Act}$ tal que $[\![\varphi]\!]^{\mathfrak{M}_E} \subseteq \mathrm{SE}^{\mathfrak{M}_E}(\pi)$ y $(\mathrm{R}_E)_\pi([\![\varphi]\!]^{\mathfrak{M}_E}) \subseteq [\![\psi]\!]^{\mathfrak{M}_E}$. Por la definición de $[E]$, $[\![\varphi]\!]^{\mathfrak{M}_E} = [\![[E]\varphi]\!]^{\mathfrak{M}}$ y $[\![\psi]\!]^{\mathfrak{M}_E} = [\![[E]\psi]\!]^{\mathfrak{M}}$. Por lo tanto, $\mathfrak{M}_E, w \models \mathsf{Kh}_i(\varphi, \psi)$ sii existe $\pi \in \mathrm{U}(i)$ con $\pi \subseteq \mathsf{Act}$ tal que $[\![[E]\varphi]\!]^{\mathfrak{M}} \subseteq \mathrm{SE}^{\mathfrak{M}_E}(\pi)$ y $(\mathrm{R}_E)_\pi([\![[E]\varphi]\!]^{\mathfrak{M}}) \subseteq [\![[E]\psi]\!]^{\mathfrak{M}}$. Sea $a \in \pi$ y $w \in \mathrm{S}_E$. Si $w \in [\![[E]\varphi]\!]^{\mathfrak{M}}$, entonces:

  - $(\mathrm{R}_E)_a(w) \neq \emptyset$ (dado que $w \in \mathrm{SE}^{\mathfrak{M}_E}(a)$) y
  - $(\mathrm{R}_E)_a(w) \subseteq [\![[E]\psi]\!]^{\mathfrak{M}}$.

  Usando la Definición 8.7, el primer ítem es equivalente a que $w \in [\![\theta]\!]^{\mathfrak{M}}$, $\mathrm{R}_a(w) \subseteq [\![\theta']\!]^{\mathfrak{M}}$ y $\mathrm{R}_a(w) \neq \emptyset$. Con esto, $(\mathrm{R}_E)_a(w) = \mathrm{R}_a(w)$ que se aplica al segundo ítem. Luego si $w \in [\![[E]\varphi]\!]^{\mathfrak{M}}$, entonces:

  - $w \in [\![\theta]\!]^{\mathfrak{M}}$,
  - $\mathrm{R}_a(w) \neq \emptyset$,
  - $\mathrm{R}_a(w) \subseteq [\![\theta']\!]^{\mathfrak{M}}$ y $\mathrm{R}_a(w) \subseteq [\![[E]\psi]\!]^{\mathfrak{M}}$.

  Como el primer ítem es independiente de π, se puede poner fuera de la expresión. Con algunos cambios se tiene que $\mathfrak{M}, w \models \mathsf{A}([E]\varphi \to \theta)$ y que si $w \in [\![[E]\varphi]\!]^{\mathfrak{M}}$, entonces:

  - $\mathrm{R}_a(w) \neq \emptyset$ (por lo tanto, $w \in \mathrm{SE}^{\mathfrak{M}}(a)$),
  - $\mathrm{R}_a(w) \subseteq [\![\theta' \wedge [E]\psi]\!]^{\mathfrak{M}}$.

  Como se demostró para $a \in \pi$ y $w \in [\![[E]\varphi]\!]^{\mathfrak{M}}$ arbitrarios, el resultado vale para todo $a \in \pi$ y $w \in [\![[E]\varphi]\!]^{\mathfrak{M}}$. Luego, $\mathfrak{M}_E, w \models \mathsf{Kh}_i(\varphi, \psi)$ sii $\mathfrak{M}, w \models \mathsf{A}([E]\varphi \to \theta)$ y existe $\pi \in \mathrm{U}(i)$ con $\pi \subseteq \mathsf{Act}$ tal que $[\![[E]\varphi]\!]^{\mathfrak{M}} \subseteq \mathrm{SE}(\pi)$ y $\mathrm{R}_\pi([E]\varphi) \subseteq [\![\theta' \wedge [E]\psi]\!]^{\mathfrak{M}}$. Esto es equivalente a que se cumpla $\mathfrak{M}, w \models (\mathsf{A}([E]\varphi \to \theta) \wedge \mathsf{Kh}_i([E]\varphi, \theta' \wedge [E]\psi))$ con $E = (\theta, \theta')$.

$\square$

Como consecuencia del Lema 8.3, dada una fórmula de $\mathsf{AUL}_{\mathsf{Kh}_i}$, se pueden eliminar las modalidades $[E]$, obteniendo así una fórmula equivalente de $\mathsf{L}_{\mathsf{Kh}_i}$. Por lo tanto, $\mathsf{AUL}_{\mathsf{Kh}_i}$ y $\mathsf{L}_{\mathsf{Kh}_i}$ son equivalentes en los modelos de $\mathbf{M_{BA}}$.

Corolario 8.3. $\mathsf{L}_{\mathsf{Kh}_i} \equiv_{\mathbf{M_{BA}}} \mathsf{AUL}_{\mathsf{Kh}_i}$.

Se tiene además que el sistema axiomático $\mathcal{L}^{\mathrm{LTS}^U}_{\mathsf{Kh}_i}$ más los axiomas de reducción son correctos y fuertemente completos con respecto a la clase $\mathbf{M_{BA}}$.

Teorema 8.2. $\mathcal{L}^{\mathrm{LTS}^U}_{\mathsf{Kh}_i}$ *más los axiomas de reducción para* $[E]$ *en la Tabla 8.2 son correctos y fuertemente completos para* $\mathsf{AUL}_{\mathsf{Kh}_i}$ *con respecto a la clase de modelos* $\mathbf{M_{BA}}$.

Demostración. Correctitud es directa del Teorema 5.1 y el Lema 8.3. Como $\mathsf{AUL}_{\mathsf{Kh}_i}$ es reducible a $\mathsf{L}_{\mathsf{Kh}_i}$ en $\mathbf{M_{BA}}$, sea Ω un conjunto consistente en $\mathsf{AUL}_{\mathsf{Kh}_i}$, Ω es equivalente a otro conjunto consistente Γ' en $\mathsf{L}_{\mathsf{Kh}_i}$ para $\mathbf{M_{BA}}$. Usando la prueba del Teorema 5.1, siendo Γ la extensión maximal consistente de Γ', Γ' es satisfacible en el modelo canónico $\mathfrak{M}^{\Gamma}$. Por la Definición 5.3, $\mathfrak{M}^{\Gamma}$ es un modelo de la clase $\mathbf{M_{BA}}$ y por lo tanto Ω es satisfacible. Con esto se prueba que todo conjunto consistente Ω en $\mathsf{AUL}_{\mathsf{Kh}_i}$ es satisfacible por un modelo de la clase $\mathbf{M_{BA}}$. $\square$

Para terminar esta sección, se tiene que el problema de satisfacibilidad para $\mathsf{AUL}_{\mathsf{Kh}_i}$ es decidible al menos en $\mathbf{M_{BA}}$. Esto quiere decir que, dada una fórmula φ de $\mathsf{AUL}_{\mathsf{Kh}_i}$, existe un procedimiento que en una cantidad finita de pasos determina si la misma es satisfacible o no en $\mathbf{M_{BA}}$.

Corolario 8.4. *El problema de satisfacibilidad para* $\mathsf{AUL}_{\mathsf{Kh}_i}$ *sobre* $\mathbf{M_{BA}}$ *es decidible.*

Demostración. Dada una fórmula φ de $\mathsf{AUL}_{\mathsf{Kh}_i}$ arbitraria, la misma se la puede traducir en una cantidad finita de pasos a una fórmula equivalente φ' de $\mathsf{L}_{\mathsf{Kh}_i}$ en $\mathbf{M_{BA}}$ (Corolario 8.3). Como el problema de satisfacibilidad para $\mathsf{L}_{\mathsf{Kh}_i}$ es decidible (Teorema 7.2), se tiene un procedimiento, si bien no determinístico, que en una cantidad finita de pasos determina si φ' es satisfacible o no. Luego de dicho procedimiento, se tienen dos resultados posibles: ***(1)*** Si φ' es satisfacible, entonces lo es en $\mathbf{M^F_{NU}}$ (Corolario 7.2). Como $\mathbf{M^F_{NU}} \subseteq \mathbf{M_{BA}}$, dado que cada $\mathrm{U}(i)$ consta de conjuntos unitarios de acciones, se tiene que φ es satisfacible en $\mathbf{M_{BA}}$. ***(2)*** Si φ' no es satisfacible, entonces tampoco lo es en $\mathbf{M_{BA}}$. Por el Corolario 8.3, φ no es satisfacible en $\mathbf{M_{BA}}$. Por lo tanto, se tiene un procedimiento tal que resuelve el problema de satisfacibilidad para $\mathsf{AUL}_{\mathsf{Kh}_i}$ en $\mathbf{M_{BA}}$. $\square$

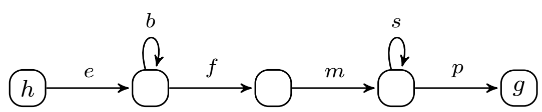
8.2. Operadores epistémicos

8.2.1. Removiendo incertidumbre entre dos planes

Dado un modelo LTS^U, la componente epistémica de un agente i, $\mathrm{U}(i)$, no sólo define los planes de los cuales el agente es 'consciente', sino también el nivel de discernimiento que puede tener entre estos. Por lo tanto, se pueden representar los cambios epistémicos en el agente definiendo operadores que modifiquen a $\mathrm{U}(i)$, en vez de actualizar el conjunto de estados S o la relación de accesibilidad. En el siguiente ejemplo ilustraremos un tipo de pasaje de información sobre un LTS^U que modifica la incertidumbre de un agente, y los efectos epistémicos del mismo.

Ejemplo 8.5. Sea $\mathfrak{M}$ el LTS^U del Ejemplo 8.1:

$\mathfrak{M}$:



$\mathrm{U}(i) = \{\{ebfmsp\}\}$
$\mathrm{U}(j) = \{\{ebfmsp, ebmfsp\}\}$

Se tiene que $\mathfrak{M} \not\models \mathsf{Kh}_j(h, g)$ dado que el plan *ebmfsp* no llega al objetivo de hacer una buena torta. Sin embargo, si el agente j pudiera ser capaz de distinguir *ebmfsp* de *ebfmsp* (que es el plan que cumple el objetivo), sería capaz de saber cómo hacer una buena torta, dado que tenga los ingredientes. Si el agente j *aprende* que el orden de las acciones importa (y que con eso *ebmfsp* es distinto de *ebfmsp*), el conjunto $\pi = \{ebfmsp, ebmfsp\}$ se dividiría en dos conjuntos unitarios. Después de esa división, el agente sabría cómo conseguir g dado h. $\square$

Con esta idea, se introduce un operador que elimina incertidumbre entre planes específicos. El operador que introduciremos está inspirado en el operador de anuncios públicos [109], excepto que la eliminación de incertidumbre se produce al nivel de planes. Esto tiene sentido ya que la componente epistemica de un modelo está constituida por planes.

En un LTS^U, existen diferentes formas de hacer distinguible dos planes que antes no lo eran. Es decir, hay diferentes formas de separar o "refinar" el conjunto que contenga a ambos planes. Para ello, daremos algunas definiciones.

Definición 8.9. Sean $\pi, \pi_1, \pi_2 \in 2^{\mathsf{Act}^*}$, y $U \subseteq 2^{\mathsf{Act}^*}$. Se escribe $\pi = \pi_1 \uplus \pi_2$ si y sólo si $\pi = \pi_1 \cup \pi_2$ y $\pi_1 \cap \pi_2 = \emptyset$. Sean $\pi \in U$ y $\pi = \pi_1 \uplus \pi_2$, se define $U^{\pi}_{\{\pi_1,\pi_2\}} \subseteq 2^{\mathsf{Act}^*}$ como el resultado de refinar π a través de $\{\pi_1, \pi_2\}$, es decir: $U^{\pi}_{\{\pi_1,\pi_2\}} := (U \setminus \{\pi\}) \cup \{\pi_1, \pi_2\}$.

Definición 8.10. Sean $U, U' \subseteq 2^{\mathsf{Act}^*}$, y $\sigma_1, \sigma_2 \in \mathsf{Act}^*$ tales que $\sigma_1 \neq \sigma_2$. Se escribe $U \rightsquigarrow^{\sigma_1}_{\sigma_2} U'$ si y sólo si

- $U' = U$ y no existe $\pi \in U$ tal que $\{\sigma_1, \sigma_2\} \subseteq \pi$, ó
- $U' = U^{\pi}_{\{\pi_1, \pi_2\}}$ para algún $\pi \in U$ tal que $\{\sigma_1, \sigma_2\} \subseteq \pi$, y $\pi_1, \pi_2 \in 2^{\mathsf{Act}^*}$ con $\pi = \pi_1 \uplus \pi_2$ y $\sigma_1 \in \pi_1$, $\sigma_2 \in \pi_2$.

Notar que la relación $\rightsquigarrow^{\sigma_1}_{\sigma_2}$ es serial. Es decir que para todo conjunto $U \subseteq 2^{\mathsf{Act}^*}$ siempre existe otro $U' \subseteq 2^{\mathsf{Act}^*}$ tal que $U \rightsquigarrow^{\sigma_1}_{\sigma_2} U'$. Más aún, si U es el conjunto de conjuntos de planes para un agente i dado en algún LTS^U (es decir, $U = \mathrm{U}(i)$) y U' cumple que $U \rightsquigarrow^{\sigma_1}_{\sigma_2} U'$, entonces la estructura resultante de reemplazar U por U' es un LTS^U.

Definición 8.11. Sean $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \mathrm{U}, \mathrm{V} \rangle$ un LTS^U y $\mathrm{U}' = \{\mathrm{U}'(i)\}_{i \in \mathsf{Agt}}$ con $\mathrm{U}'(i) \subseteq 2^{\mathsf{Act}^*}$. Sean $\sigma_1, \sigma_2 \in \mathsf{Act}^*$. Se escribe $\mathrm{U} \rightsquigarrow^{\sigma_1}_{\sigma_2} \mathrm{U}'$ sii para cada $i \in \mathsf{Agt}$, $\mathrm{U}(i) \rightsquigarrow^{\sigma_1}_{\sigma_2} \mathrm{U}'(i)$. Se denota $\mathfrak{M}^{\mathrm{U}}_{\mathrm{U}'} = \langle \mathrm{S}, \mathrm{R}, \mathrm{U}', \mathrm{V} \rangle$ el LTS^U obtenido de reemplazar U por U′.

Con estas herramientas, se introduce la nueva modalidad $\langle \sigma_1 \not\sim \sigma_2 \rangle$, interpretada como una acción la cual todos los agentes que son conscientes de los planes σ_1 y σ_2, aprenden que estos son diferentes. A continuación se introduce la sintáxis de esta extensión de $\mathsf{L}_{\mathsf{Kh}_i}$.

Definición 8.12. Las fórmulas del lenguaje $\mathsf{L}_{\mathsf{Ref}}$ ("refinamiento") están dadas por la siguiente gramática:

$$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid \mathsf{Kh}_i(\varphi, \varphi) \mid \langle \sigma_1 \not\sim \sigma_2 \rangle \varphi,$$

con $p \in \mathsf{Prop}$, $i \in \mathsf{Agt}$ y $\sigma_1, \sigma_2 \in \mathsf{Act}^*$ tales que $\sigma_1 \neq \sigma_2$. Las constantes y demás conectores booleanos $\bot$, $\top$, $\wedge$, $\rightarrow$, y $\leftrightarrow$ se definen de la forma usual. Se define la modalidad dual $[\sigma_1 \not\sim \sigma_2]\varphi := \neg\langle \sigma_1 \not\sim \sigma_2 \rangle \neg\varphi$. Las fórmulas de la forma $\langle \sigma_1 \not\sim \sigma_2 \rangle \varphi$ se leen como *"después de que se establece que σ_1 y σ_2 son distinguibles, φ se cumple"*.

Con esto, presentamos la semántica para la modalidad $\langle \sigma_1 \not\sim \sigma_2 \rangle$ utilizando las definiciones anteriores.

Definición 8.13. Sean $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \mathrm{U}, \mathrm{V} \rangle$ un LTS^U y $w \in \mathrm{S}$. Para $\sigma_1 \neq \sigma_2$,

$$\mathfrak{M}, w \models \langle \sigma_1 \not\sim \sigma_2 \rangle \varphi \text{ sii existe U}' \text{ tal que U} \rightsquigarrow^{\sigma_1}_{\sigma_2} \mathrm{U}' \text{ y } \mathfrak{M}^{\mathrm{U}}_{\mathrm{U}'}, w \models \varphi.$$

Tomando el Ejemplo 8.5, la fórmula $\langle ebmfsp \not\sim ebfmsp \rangle \mathsf{Kh}_j(h, g)$ se lee como *"después de que se establece que ebmfsp y ebfmsp son planes distinguibles, el agente j sabe cómo hacer una buena torta, dado que tenga los ingredientes"*. Esta modalidad además es normal y serial.

Proposición 8.6. *Sean $\sigma_1, \sigma_2 \in \mathsf{Act}^*$, ψ y φ fórmulas arbitrarias de $\mathsf{L}_{\mathsf{Ref}}$, las siguientes propiedades se cumplen:*

1. $\models [\sigma_1 \not\sim \sigma_2](\varphi \rightarrow \psi) \rightarrow ([\sigma_1 \not\sim \sigma_2]\varphi \rightarrow [\sigma_1 \not\sim \sigma_2]\psi)$.
2. *Si* $\models \varphi$, *entonces* $\models [\sigma_1 \not\sim \sigma_2]\varphi$.
3. $\models [\sigma_1 \not\sim \sigma_2]\varphi \rightarrow \langle \sigma_1 \not\sim \sigma_2 \rangle \varphi$.

Demostración. El Ítem 1 se sigue del patrón estándar de $\forall$ en la semántica de $[\sigma_1 \not\sim \sigma_2]$. El Ítem 2 se cumple porque la estructura resultante del operador es un LTS^U. Por otro lado, la serialidad de $\rightsquigarrow^{\sigma_1}_{\sigma_2}$ garantiza el Ítem 3. □

Por consiguiente, para un tipo específico de fórmulas, la modalidad dinámica preserva el conocimiento y hasta puede generar nuevo.

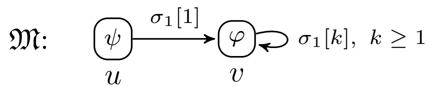
Proposición 8.7. *Sean $\sigma_1, \sigma_2 \in \mathsf{Act}^*$ y ψ, φ fórmulas* proposicionales.

1. $\models \mathsf{Kh}_i(\psi, \varphi) \rightarrow [\sigma_1 \not\sim \sigma_2]\mathsf{Kh}_i(\psi, \varphi)$.
2. *Si ψ y φ son satisfacibles, entonces $\neg\mathsf{Kh}_i(\psi, \varphi) \wedge \langle \sigma_1 \not\sim \sigma_2 \rangle \mathsf{Kh}_i(\psi, \varphi)$ es satisfacible.*

Demostración. Para el Ítem 1, sea $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \mathrm{U}, \mathrm{V} \rangle$, si $\mathfrak{M}, w \models \mathsf{Kh}_i(\psi, \varphi)$, entonces existe $\pi \in \mathrm{U}(i)$ tal que $[\![\psi]\!]^{\mathfrak{M}} \subseteq \mathrm{SE}(\pi)$ y $\mathrm{R}_\pi([\![\psi]\!]^{\mathfrak{M}}) \subseteq [\![\varphi]\!]^{\mathfrak{M}}$. Sean $\sigma_1, \sigma_2 \in \mathsf{Act}^*$ y U' tal que $\mathrm{U} \rightsquigarrow^{\sigma_1}_{\sigma_2} \mathrm{U}'$. Como ψ y φ son proposicionales, $[\![\psi]\!]^{\mathfrak{M}} = [\![\psi]\!]^{\mathfrak{M}^{\mathrm{U}}_{\mathrm{U}'}}$ y $[\![\varphi]\!]^{\mathfrak{M}} = [\![\varphi]\!]^{\mathfrak{M}^{\mathrm{U}}_{\mathrm{U}'}}$. Si $\sigma_1 \notin \pi$ ó $\sigma_2 \notin \pi$, entonces $\pi \in \mathrm{U}'$ y sigue siendo el testigo para $\mathsf{Kh}_i(\psi, \varphi)$ en $\mathfrak{M}^{\mathrm{U}}_{\mathrm{U}'}$. Si $\sigma_1, \sigma_2 \in \pi$, sea $\{\pi_1, \pi_2\}$ la partición de π tal que $\mathrm{U}(i) \rightsquigarrow^{\sigma_1}_{\sigma_2} \mathrm{U}(i)^{\pi}_{\{\pi_1, \pi_2\}}$. Para $k \in \{1, 2\}$, como $\pi_k \subseteq \pi$, se tiene que para todo $w \in \mathrm{D}_{\mathfrak{M}}$ tal que π es fuertemente ejecutable en w, implica que π_k es fuertemente ejecutable en w. Por lo tanto, $\mathrm{SE}(\pi) \subseteq \mathrm{SE}(\pi_k)$. Además, $\mathrm{R}_{\pi_k} \subseteq \mathrm{R}_\pi$ dado que para todo $(w, v) \in \mathrm{R}_{\pi_k}$, como $\pi_k \subseteq \pi$, se tiene que $(w, v) \in \mathrm{R}_\pi$. Por lo tanto, $[\![\psi]\!]^{\mathfrak{M}^{\mathrm{U}}_{\mathrm{U}'}} \subseteq \mathrm{SE}(\pi) \subseteq \mathrm{SE}(\pi_k)$ y $\mathrm{R}_{\pi_k}([\![\psi]\!]^{\mathfrak{M}^{\mathrm{U}}_{\mathrm{U}'}}) \subseteq \mathrm{R}_\pi([\![\psi]\!]^{\mathfrak{M}^{\mathrm{U}}_{\mathrm{U}'}}) \subseteq [\![\varphi]\!]^{\mathfrak{M}^{\mathrm{U}}_{\mathrm{U}'}}$, para $k \in \{1, 2\}$. Ello quiere decir que, como el agente i sabía cómo ir de estados ψ a estados φ vía π, debilitar π haciendo una partición no modifica la propiedad de testigo que tenía antes. Esto habilita a que el agente pueda elegir entre π_1 ó π_2 como testigo. Con esto, todos los casos para σ_1 y σ_2 con π están cubiertos y se tiene que $\mathfrak{M}^{\mathrm{U}}_{\mathrm{U}'}, w \models \mathsf{Kh}_i(\psi, \varphi)$ para todo U' tal que $\mathrm{U} \rightsquigarrow^{\sigma_1}_{\sigma_2} \mathrm{U}'$. Por lo tanto, $\mathfrak{M}^{\mathrm{U}}_{\mathrm{U}'}, w \models [\sigma_1 \not\sim \sigma_2]\mathsf{Kh}_i(\psi, \varphi)$

Para el Ítem 2 se define el modelo $\mathfrak{M} = \langle \{u, v\}, \mathrm{R}, \{\mathrm{U}(i)\}, \mathrm{V} \rangle$ tal que

- $\mathfrak{M}, u \models \psi$ y $\mathfrak{M}, v \models \varphi$;
- $\mathrm{R}_{\sigma_1[1]} = \{(u, v), (v, v)\}$ y $\mathrm{R}_{\sigma_1[j]} = \{(v, v)\}$ para todo $j = 2, \ldots, |\sigma_1|$;
- $\mathrm{R}_{\sigma_2[k]} = \emptyset$ para algún $k = 1, \ldots, |\sigma_2|$ (pues $\sigma_1 \neq \sigma_2$, se omite σ_2 en la gráfica); y
- $\mathrm{U}(i) = \{\{\sigma_1, \sigma_2\}\}$.



Como $u \notin \mathrm{SE}(\sigma_2) = \mathrm{SE}(\{\sigma_1, \sigma_2\})$, $\mathfrak{M} \models \neg\mathsf{Kh}_i(\psi, \varphi)$ pero $\mathfrak{M} \models \langle \sigma_1 \not\sim \sigma_2 \rangle \mathsf{Kh}_i(\psi, \varphi)$ dado que $\{\{\sigma_1, \sigma_2\}\} \rightsquigarrow^{\sigma_1}_{\sigma_2} \{\{\sigma_1\}, \{\sigma_2\}\}$ y $\{\sigma_1\}$ puede elegirse como testigo. □

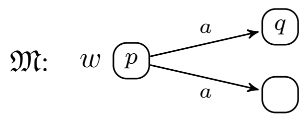
Sin embargo esto no se cumple para fórmulas arbitrarias de $\mathsf{L}_{\mathsf{Ref}}$. Al considerar las modalidades Kh_i y $\langle \sigma_1 \not\sim \sigma_2 \rangle$, no es posible generalizar estas propiedades.

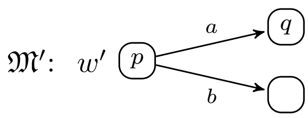
Ejemplo 8.6. Teniendo en cuenta el Ejemplo 8.1, sea $\theta = \mathsf{Kh}_j(h, g)$, $\mathfrak{M} \not\models \mathsf{Kh}_j(\theta, \neg\theta) \rightarrow [\mathit{ebmfsp} \not\sim \mathit{ebfmsp}]\mathsf{Kh}_j(\theta, \neg\theta)$. Por otro lado, $\neg\mathsf{Kh}_i(\top, \bot) \wedge \langle \sigma_1 \not\sim \sigma_2 \rangle \mathsf{Kh}_i(\top, \bot)$ no es satisfacible. □

Como podemos notar, la nueva modalidad puede hablar explícitamente sobre los planes específicos, a diferencia del operador Kh_i. Dados dos planes, al declarar en un modelo que ambos son distinguibles entre sí, los efectos en el conocimiento de un agente i varían dependiendo de su conocimiento previo al anuncio: Si i no era consciente de al menos uno de los dos planes, o era consciente de ambos, pero ya los distinguía entre sí, entonces el anuncio no cambia su percepción. Sin embargo, si i era consciente de ambos pero no era capaz de distinguirlos, gracias al anuncio, su conocimiento cambia. La siguiente proposición considera estas dos alternativas y se las utiliza para mostrar que $\mathsf{L}_{\mathsf{Ref}}$ es más expresiva que $\mathsf{L}_{\mathsf{Kh}_i}$.

Proposición 8.8. $\mathsf{L}_{\mathsf{Kh}_i} \prec \mathsf{L}_{\mathsf{Ref}}$.

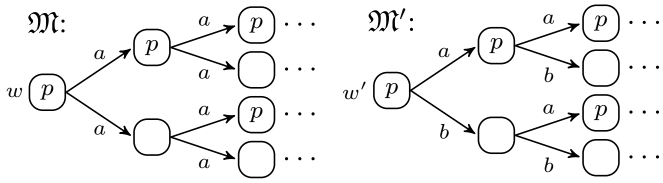
Demostración. Como $\mathsf{L}_{\mathsf{Ref}}$ es una extensión de $\mathsf{L}_{\mathsf{Kh}_i}$, se tiene que $\mathsf{L}_{\mathsf{Kh}_i} \preceq \mathsf{L}_{\mathsf{Ref}}$. Para probar que $\mathsf{L}_{\mathsf{Ref}} \not\preceq \mathsf{L}_{\mathsf{Kh}_i}$, se deben mostrar dos modelos LTS^Us bisimilares en $\mathsf{L}_{\mathsf{Kh}_i}$ y una fórmula de $\mathsf{L}_{\mathsf{Ref}}$ que los distingan. Sean $\mathfrak{M}$ y $\mathfrak{M}'$ dos modelos de un agente, con $\mathrm{U}(i) := \{\{a\}\}$ y $\mathrm{U}'(i) := \{\{a, b\}\}$, respectivamente:





Los modelos son bisimilares en $\mathsf{L}_{\mathsf{Kh}_i}$ ($\mathfrak{M}, w \leftrightarroweq_{\mathsf{L}_{\mathsf{Kh}_i}} \mathfrak{M}', w'$), y por el Teorema 6.1, $\mathfrak{M}, w \leftrightsquigarrow_{\mathsf{L}_{\mathsf{Kh}_i}} \mathfrak{M}', w'$, es decir, satisfacen las mismas fórmulas en $\mathsf{L}_{\mathsf{Kh}_i}$. Particularmente, ambos satisfacen $\neg\mathsf{Kh}_i(p, q)$. Sin embargo, $\mathfrak{M}, w \not\models \langle a \not\sim b \rangle \mathsf{Kh}_i(p, q)$ dado que $\mathrm{U}(i) \rightsquigarrow^{a}_{b} \mathrm{U}(i)$ únicamente, mientras que $\mathfrak{M}', w' \models \langle a \not\sim b \rangle \mathsf{Kh}_i(p, q)$, dado que existe $\mathrm{U}''(i) = \{\{a\}, \{b\}\}$ tal que $\mathrm{U}'(i) \rightsquigarrow^{a}_{b} \mathrm{U}''(i)$ y $\mathfrak{M}^{\mathrm{U}'}_{\mathrm{U}''}, w \models \mathsf{Kh}_i(p, q)$. Con esto, $\mathsf{L}_{\mathsf{Ref}} \not\preceq \mathsf{L}_{\mathsf{Kh}_i}$ y por ende, $\mathsf{L}_{\mathsf{Kh}_i} \prec \mathsf{L}_{\mathsf{Ref}}$. □

Ejemplo 8.7. Otro ejemplo interesante sobre la expresividad de $\mathsf{L}_{\mathsf{Ref}}$ es el siguiente. Consideremos los siguientes árboles binarios, cada uno dividiéndose en p y $\neg p$. Consideremos además que hay un sólo agente i y que los conjuntos de planes respectivos son $\mathrm{U}(i) = \{\{a\}\}$ y $\mathrm{U}'(i) = \{\{a, b\}\}$.

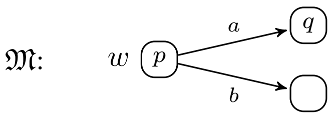


Se tiene que $\mathfrak{M}, w \;\underline{\leftrightarrow}_{\mathsf{L}_{\mathsf{Kh}_i}}\; \mathfrak{M}', w'$ (y por lo tanto $\mathfrak{M}, w \leftrightsquigarrow_{\mathsf{L}_{\mathsf{Kh}_i}} \mathfrak{M}', w'$), pero de manera similar $\mathfrak{M}, w \not\models \langle a \not\sim b\rangle \mathsf{Kh}_i(p, p)$ mientras que $\mathfrak{M}', w' \models \langle a \not\sim b\rangle \mathsf{Kh}_i(p, p)$. $\square$

Tal como hemos explicado al principio de este capítulo, nos interesa contar con sistemas axiomáticos que nos permitan caracterizar los teoremas de nuestras lógicas dinámicas. Sin embargo, usualmente esto no resulta ser fácil [7]. En particular, y para concluir esta subsección, la regla de sustitución uniforme, que suele ser un indicativo de buen comportamiento a la hora de axiomatizar, no es admisible para $\mathsf{L}_{\mathsf{Ref}}$. La misma establece que, para toda fórmula válida φ de $\mathsf{L}_{\mathsf{Ref}}$ ($\models \varphi$), cualquier reemplazo sintáctico simultáneo de alguna de sus variables proposicionales (p) por una fórmula arbitraria (ψ) preserva su validez ($\models \varphi[p := \psi]$). Este no cumplimiento se debe a que, como la modalidad $\langle \sigma_1 \not\sim \sigma_2\rangle$ habla sobre planes, el reemplazo sintáctico de una fórmula válida, con una ocurrencia de $\langle \sigma_1 \not\sim \sigma_2\rangle$, por una fórmula $\mathsf{Kh}_i(\psi, \varphi)$ no necesariamente se cumple en todos los modelos.

Proposición 8.9. *Sustitución uniforme no se cumple en* $\mathsf{L}_{\mathsf{Ref}}$.

Demostración. Debido a que la modalidad induce un cambio epistémico, las variables proposicionales no se ven afectadas en ningún momento. Por lo tanto, la fórmula $p \to [a \not\sim b]p$ es válida para cualquier $a, b \in \mathsf{Act}$. Sea $\mathfrak{M}$ el siguiente modelo tal que $\mathrm{U}(i) := \{\{a, b\}\}$:



Si se sustituye p por $\neg\mathsf{Kh}_i(p, q)$, la fórmula resultante es $\neg\mathsf{Kh}_i(p, q) \to [a \not\sim b]\neg\mathsf{Kh}_i(p, q)$. Está claro que $\mathfrak{M} \models \neg\mathsf{Kh}_i(p, q)$, pero $\mathfrak{M} \not\models [a \not\sim b]\neg\mathsf{Kh}_i(p, q)$ pues existe $\mathrm{U}'(i) = \{\{a\}, \{b\}\}$ tal que $\{\{a, b\}\} \rightsquigarrow^a_b \mathrm{U}'(i)$ y $\mathfrak{M}^{\mathrm{U}}_{\mathrm{U}'} \models \mathsf{Kh}_i(p, q)$. Por lo tanto, $\neg\mathsf{Kh}_i(p, q) \to [a \not\sim b]\neg\mathsf{Kh}_i(p, q)$ no es válida. $\square$

8.2.2. Refinamiento arbitrario sobre planes

Como se mencionó anteriormente, la operación $\langle \sigma_1 \not\sim \sigma_2\rangle$ puede ser vista como una forma particular de remover incertidumbre: similar a un anuncio público, indica precisamente que dos planes son distinguibles y a su vez cuantifica sobre las diferentes formas de distinguirlas. Sin embargo, podemos pensar en una manera más abstracta de refinar la incertidumbre: tomando inspiración de otras propuestas que cuantifican sobre acciones epistémicas (por ejemplo, los anuncios arbitrarios [14], actualizaciones arbitrarias de relaciones [110], anuncios de grupo [1] y anuncios de coalición [2]), se cuantifica sobre todas las formas posibles en las que la incertidumbre del agente puede ser refinada. Procedemos a introducir la sintáxis de esta extensión de $\mathsf{L}_{\mathsf{Kh}_i}$.

Definición 8.14. Las fórmulas del lenguaje $\mathsf{L}_{\mathsf{ARef}}$ ("refinamiento arbitrario") están dadas por la siguiente gramática:

$$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid \mathsf{Kh}_i(\varphi, \varphi) \mid \langle \not\sim \rangle\varphi,$$

con $p \in \mathsf{Prop}$ y $i \in \mathsf{Agt}$. Las constantes y demás conectores booleanos $\bot$, $\top$, $\wedge$, $\to$, y $\leftrightarrow$ se definen de la forma usual. Se define la modalidad dual $[\not\sim]\varphi := \neg\langle \not\sim \rangle\neg\varphi$. Las fórmulas de la forma $\langle \not\sim \rangle\varphi$ se leen como *"después de que se establece que dos planes son distinguibles, φ se cumple"*.

Luego, se tiene la semántica para la modalidad $\langle \not\sim \rangle$.

Definición 8.15. Sea $\mathfrak{M}$ un LTS^U y $w \in \mathrm{D}_{\mathfrak{M}}$. Entonces,

$$\mathfrak{M}, w \models \langle \not\sim \rangle \varphi \text{ sii existen } \sigma_1, \sigma_2 \in \mathsf{Act}^* \text{ distintos tales que } \mathfrak{M}, w \models \langle \sigma_1 \not\sim \sigma_2 \rangle \varphi.$$

Como sucedía con la modalidad $\langle \sigma_1 \not\sim \sigma_2 \rangle$, la modalidad $\langle \not\sim \rangle$ es normal y serial.

Proposición 8.10. *Sean ψ y φ fórmulas arbitrarias de $\mathsf{L}_{\mathsf{ARef}}$, las siguientes propiedades se cumplen:*

1. $\models [\not\sim](\varphi \to \psi) \to ([\not\sim]\varphi \to [\not\sim]\psi)$.
2. *Si* $\models \varphi$, *entonces* $\models [\not\sim]\varphi$.
3. $\models [\not\sim]\varphi \to \langle \not\sim \rangle \varphi$.

Demostración. Sea $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \mathrm{U}, \mathrm{V} \rangle$ y $w \in \mathrm{S}$:

1. Supongamos que $\mathfrak{M}, w \models [\not\sim](\varphi \to \psi)$ y $\mathfrak{M}, w \models [\not\sim]\varphi$. Sean $\sigma_1, \sigma_2 \in \mathsf{Act}^*$ planes arbitrarios. Entonces $\mathfrak{M}, w \models [\sigma_1 \not\sim \sigma_2](\varphi \to \psi)$ y $\mathfrak{M}, w \models [\sigma_1 \not\sim \sigma_2]\varphi$. Usando el Ítem 1 de la Proposición 8.6, $\mathfrak{M}, w \models [\sigma_1 \not\sim \sigma_2]\psi$. Por lo tanto, $\mathfrak{M}, w \models [\not\sim]\psi$.

2. Supongamos que $\models \varphi$. Sean $\sigma_1, \sigma_2 \in \mathsf{Act}^*$ planes arbitrarios. Usando el Ítem 2 de la Proposición 8.6, $\models [\sigma_1 \not\sim \sigma_2]\varphi$. Por lo tanto, $\models [\not\sim]\varphi$.

3. Supongamos que $\mathfrak{M}, w \models [\not\sim]\varphi$. Entonces para todo $\sigma_1, \sigma_2 \in \mathsf{Act}^*$, $\mathfrak{M}, w \models [\sigma_1 \not\sim \sigma_2]\varphi$. Como Act no es vacío, se eligen $\alpha_1, \alpha_2 \in \mathsf{Act}^*$, y se tiene que $\mathfrak{M}, w \models [\alpha_1 \not\sim \alpha_2]\varphi$. Usando el Ítem 3 de la Proposición 8.6, $\mathfrak{M}, w \models \langle \alpha_1 \not\sim \alpha_2 \rangle \varphi$. Por lo tanto, $\mathfrak{M}, w \models \langle \not\sim \rangle \varphi$.

□

De manera análoga, la modalidad dinámica preserva el conocimiento y hasta puede generar nuevo para fórmulas proposicionales.

Proposición 8.11. *Sean ψ, φ fórmulas* proposicionales.

1. $\models \mathsf{Kh}_i(\psi, \varphi) \to [\not\sim]\mathsf{Kh}_i(\psi, \varphi)$.
2. *Si ψ y φ son satisfacibles, entonces* $\neg\mathsf{Kh}_i(\psi, \varphi) \wedge \langle \not\sim \rangle \mathsf{Kh}_i(\psi, \varphi)$ *es satisfacible.*

Demostración. Similar a la vista en la Proposición 8.7. □

Además satisface las propiedades de monotonía y debilitamiento. Pero no cumple las versiones dinámicas de los axiomas 4 y 5 (usualmente conocidas como *introspección positiva* y *negativa*).

Proposición 8.12. *Sean φ, ψ fórmulas de $\mathsf{L}_{\mathsf{ARef}}$ arbitrarias. Se cumple lo siguiente:*

1. $\models \langle \not\sim \rangle \varphi \to \langle \not\sim \rangle (\varphi \vee \psi)$ *y* $\models [\not\sim]\varphi \to [\not\sim](\varphi \vee \psi)$ *(monotonía).*
2. $\models \langle \not\sim \rangle (\varphi \wedge \psi) \to \langle \not\sim \rangle \varphi$ *y* $\models [\not\sim](\varphi \wedge \psi) \to [\not\sim]\varphi$ *(debilitamiento).*
3. $\not\models [\not\sim]\varphi \to [\not\sim][\not\sim]\varphi$ *(introspección positiva).*
4. $\not\models \neg[\not\sim]\varphi \to [\not\sim]\neg[\not\sim]\varphi$ *(introspección negativa).*

Demostración. Usando la definición de $\models$, monotonía y debilitamiento se cumplen. Para la introspección positiva, sea $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \mathrm{U}, \mathrm{V} \rangle$ el siguiente modelo con $\mathrm{U}(i) \in \mathrm{U}$ tal que $\mathrm{U}(i) := \{\{a, b, c\}\}$.

$$\mathfrak{M}: \quad \begin{array}{c} w\ \boxed{p} \xrightarrow{\quad a \quad} \boxed{q} \\[1em] \boxed{r} \xrightarrow{\quad b \quad} \boxed{s} \end{array}$$

Por un lado, se puede ver que $\mathfrak{M}, w \models [\not\sim]\neg(\mathsf{Kh}(p, q) \wedge \mathsf{Kh}(r, s))$. Por otro lado, se tiene que $\mathfrak{M}, w \models \langle a \not\sim b \rangle \langle b \not\sim c \rangle (\mathsf{Kh}(p, q) \wedge \mathsf{Kh}(r, s))$. Por lo tanto, se tiene que $\mathfrak{M}, w \models \langle \not\sim \rangle \langle \not\sim \rangle (\mathsf{Kh}(p, q) \wedge \mathsf{Kh}(r, s))$. Para la introspección negativa, notar que $\mathfrak{M}, w \models \neg[\not\sim]\mathsf{Kh}(p, q)$ pero $\mathfrak{M}, w \models \neg[\not\sim]\neg[\not\sim]\mathsf{Kh}(p, q)$, dado que $\mathfrak{M}, w \models \langle a \not\sim b \rangle [\not\sim]\mathsf{Kh}(p, q)$. □

Usando los argumentos de la Proposición 8.8, como la modalidad $\langle\not\sim\rangle$ puede hablar sobre planes cuantificando sobre estos, se tiene que $\mathsf{L}_{\mathsf{ARef}}$ es más expresiva que $\mathsf{L}_{\mathsf{Kh}_i}$.

Proposición 8.13. $\mathsf{L}_{\mathsf{Kh}_i} \prec \mathsf{L}_{\mathsf{ARef}}$.

Demostración. Tomando el par de modelos bisimilares de la demostración de la Proposición 8.8, se tiene que para todo $\sigma_1, \sigma_2 \in \mathsf{Act}^*$, $\mathfrak{M}, w \not\models \langle \sigma_1 \not\sim \sigma_2 \rangle \mathsf{Kh}_i(p,q)$ dado que si $\mathrm{U}(i) \rightsquigarrow^{\sigma_1}_{\sigma_2} \mathrm{U}'(i)$, necesariamente $\mathrm{U}'(i) = \mathrm{U}(i)$. Por lo tanto $\mathfrak{M}, w \models [\sigma_1 \not\sim \sigma_2]\neg\mathsf{Kh}_i(p,q)$ y con esto $\mathfrak{M}, w \models [\not\sim]\neg\mathsf{Kh}_i(p,q)$

Mientras tanto, $\mathfrak{M}', w' \models \langle a \not\sim b\rangle \mathsf{Kh}_i(p,q)$, dado que existe $\mathrm{U}''(i) = \{\{a\},\{b\}\}$ tal que $\mathrm{U}'(i) \rightsquigarrow^a_b \mathrm{U}''(i)$ y $\mathfrak{M}'^{\mathrm{U}'}_{\mathrm{U}''}, w \models \mathsf{Kh}_i(p,q)$. Con esto, $\mathfrak{M}', w' \models \langle\not\sim\rangle \mathsf{Kh}_i(p,q)$ y por ende, $\mathfrak{M}', w' \not\models [\not\sim]\neg\mathsf{Kh}_i(p,q)$ □

Por último, como sucedió con $\mathsf{L}_{\mathsf{Ref}}$, la sustitución uniforme tampoco se cumple en esta lógica. Nuevamente, dado que la modalidad $\langle\not\sim\rangle$ habla sobre planes, el reemplazo sintáctico de una fórmula válida, con una ocurrencia de este operador, por una fórmula $\mathsf{Kh}_i(\psi,\varphi)$ no necesariamente se cumple en todos los modelos.

Proposición 8.14. *Sustitución uniforme no se cumple en* $\mathsf{L}_{\mathsf{ARef}}$.

Demostración. La prueba es similar a la Proposición 8.9, tomando $p \to [\not\sim]p$ como la fórmula válida original y sustituyendo p por $\neg\mathsf{Kh}_i(p,q)$. □

8.2.3. Aprender cómo orientado a objetivos

A lo largo de este trabajo, se puede notar que los operadores de saber cómo están definidos en base a *cumplir objetivos*: el agente busca un curso de acción apropiado que le haga conseguir un determinado estado. Tiene sentido definir un operador que, cuando sea posible, *garantice* que el agente *aprenda cómo* conseguir un objetivo. Esta acción puede ser entendida como un operador de aprender cómo orientado a objetivos: se busca una forma de particionar *algunos* de los conjuntos de planes π de tal forma que el agente sepa cómo conseguir φ dado ψ.

Sea L_{Lh} (del inglés "learning how") una extensión de $\mathsf{L}_{\mathsf{Kh}_i}$ con la modalidad dinámica

$$\langle\psi,\varphi\rangle_i\chi := \langle\not\sim\rangle(\mathsf{Kh}_i(\psi,\varphi) \wedge \chi),$$

siendo su 'dual' $[\psi,\varphi]_i\chi := \neg\langle\psi,\varphi\rangle_i\neg\chi$. Se define además $\mathsf{L}_i(\psi,\varphi) := \langle\psi,\varphi\rangle_i\top$, una abreviación que establece que *"el agente i puede aprender cómo cumplir φ dado ψ"*. Este nuevo operador dinámico es una modalidad ternaria que expresa que el agente es capaz de aprender cómo cumplir φ dado ψ, y que después de que esta operación de aprendizaje sucede, χ se cumple. Por otro lado, la modalidad L_i testea qué habilidades puede aprender el agente i. Notar que L_{Lh} no es más que un fragmento sintáctico de $\mathsf{L}_{\mathsf{ARef}}$ y por lo tanto es reducible a esta última. Sin embargo, como veremos en la siguiente proposición, $\mathsf{L}_{\mathsf{ARef}}$ es reducible a L_{Lh} y por ende ambas son equivalentes.

Proposición 8.15. $\mathsf{L}_{\mathsf{Lh}} \equiv \mathsf{L}_{\mathsf{ARef}}$

Demostración. Como L_{Lh} es un fragmento sintáctico de $\mathsf{L}_{\mathsf{ARef}}$, se tiene que $\mathsf{L}_{\mathsf{Lh}} \preceq \mathsf{L}_{\mathsf{ARef}}$. Para la otra reducción, se tiene la función de traducción $T : \mathsf{L}_{\mathsf{ARef}} \to \mathsf{L}_{\mathsf{Lh}}$ que se define inductivamente de la siguiente forma:

- $T(p) = p$,
- $T(\neg\varphi_1) = \neg T(\varphi_1)$,
- $T(\varphi_1 \vee \varphi_2) = T(\varphi_1) \vee T(\varphi_2)$,
- $T(\mathsf{Kh}_i(\varphi_1,\varphi_2)) = \mathsf{Kh}_i(T(\varphi_1), T(\varphi_2))$, y
- $T(\langle\not\sim\rangle\varphi_1) = \langle\not\sim\rangle(\mathsf{Kh}_i(\bot,\bot) \wedge T(\varphi_1)) = \langle\bot,\bot\rangle_i T(\varphi_1)$.

El objetivo es demostrar por inducción que para cualquier fórmula φ de $\mathsf{L}_{\mathsf{ARef}}$, existe una fórmula $\varphi' = T(\varphi)$ tal que para todo LTS^U $\mathfrak{M}$ y $w \in \mathrm{D}_{\mathfrak{M}}$, se tiene que $\mathfrak{M}, w \models \varphi$ sii $\mathfrak{M}, w \models T(\varphi)$. Con esta propiedad, $\mathsf{L}_{\mathsf{ARef}} \preceq \mathsf{L}_{\mathsf{Lh}}$. Los primeros cuatro casos se cumplen de manera directa usando la definición de $\models$ e hipótesis inductiva. Por lo tanto, nos enfocaremos en el caso de la modalidad $\langle \not\sim \rangle$. Sean $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \mathrm{U}, \mathrm{V} \rangle$ un LTS^U y $w \in \mathrm{D}_{\mathfrak{M}}$, $\mathfrak{M}, w \models \langle \not\sim \rangle \varphi_1$ si y sólo si existen $\sigma_1, \sigma_2 \in \mathsf{Act}^*$ tales que $\mathfrak{M}, w \models \langle \sigma_1 \not\sim \sigma_2 \rangle \varphi_1$. Esto equivale a que existen $\sigma_1, \sigma_2 \in \mathsf{Act}^*$ y U' tales que $\mathrm{U} \rightsquigarrow^{\sigma_1}_{\sigma_2} \mathrm{U}'$ y $\mathfrak{M}^{\mathrm{U}}_{\mathrm{U}'}, w \models \varphi_1$. Por hipótesis inductiva, $\mathfrak{M}^{\mathrm{U}}_{\mathrm{U}'}, w \models T(\varphi_1)$. Como $\mathsf{Kh}_i(\bot, \bot)$ es una fórmula válida (pues por definición $\mathrm{U}(i) \neq \emptyset$), se cumple en cualquier modelo LTS^U. Luego, $\mathfrak{M}^{\mathrm{U}}_{\mathrm{U}'}, w \models \mathsf{Kh}_i(\bot, \bot) \wedge T(\varphi_1)$ y usando las definiciones de $\models$ y T, esto equivale a que $\mathfrak{M}, w \models T(\langle \not\sim \rangle \varphi_1)$. Por lo tanto, $\mathsf{L}_{\mathsf{ARef}} \preceq \mathsf{L}_{\mathsf{Lh}}$ y por consiguiente $\mathsf{L}_{\mathsf{ARef}} \equiv \mathsf{L}_{\mathsf{Lh}}$. □

Por lo tanto, independientemente de cuál de estas lógicas se utilice, se puede definir el concepto de aprender cómo con estas. Volviendo a estos operadores dinámicos, la siguiente proposición ilustra algunas propiedades de estas modalidades.

Proposición 8.16. *Sean ψ y φ fórmulas arbitrarias de L_{Lh}, se cumple lo siguiente:*

1. $\not\models \mathsf{L}_i(\psi, \varphi)$*; y*
2. $\mathsf{L}_i(\psi, \varphi) \wedge \mathsf{L}_i(\varphi, \neg\psi)$ *es satisfacible.*

Demostración. El Ítem 1 muestra que un agente no puede aprender todas las habilidades. La (in)disponibilidad de ciertas acciones en un LTS^U restringe lo que puede ser aprendido. Teniendo en cuenta el siguiente LTS^U $\mathfrak{M}$ con un único agente, tal que $\mathrm{U}(i) = \{\{ab, a\}, \{\epsilon\}\}$.

$$\mathfrak{M}: \quad w\,\boxed{p} \xrightarrow{\quad a \quad} \boxed{p} \xrightarrow{\quad b \quad} \boxed{p, r}$$

Notar que $\mathfrak{M}, w \not\models \mathsf{Kh}_i(p, r)$ dado que $\{ab, a\}$ no es fuertemente ejecutable en todos los estados p, sólo lo es en w. Por otro lado, $\{\epsilon\}$ es fuertemente ejecutable en todos los estados pero no siempre llega a estados r. Más aún, $\mathfrak{M}, w \not\models \mathsf{L}_i(p, r)$. El conjunto $\{\epsilon\}$ no puede ser refinado, y ningún refinamiento de $\{ab, a\}$ funciona. Por lo tanto, el agente i no puede aprender cómo cumplir r dado p.

Para el Ítem 2, sea $\mathfrak{M}'$ el modelo de la Proposición 8.8. $\mathfrak{M}', w' \not\models \mathsf{Kh}_i(p, q)$ pero existe una forma de aprender cómo conseguir q dado p: es posible particionar el conjunto $\{a, b\}$ en $\{a\}$ y $\{b\}$. Por lo tanto, $\mathfrak{M}', w' \models \mathsf{L}_i(p, q)$ (con $\{a\}$ como testigo), pero también $\mathfrak{M}', w' \models \mathsf{L}_i(p, \neg q)$ (con $\{b\}$ como testigo). □

El Ítem 1 muestra cómo en ciertos escenarios no es posible aprender nada nuevo. Por ejemplo, es posible que no haya una forma de aprender cómo curar una enfermedad si no hay un médico disponible o una vacuna desarrollada. El Ítem 2 muestra cómo el agente puede ser capaz de aprender cómo no sólo hacer una fórmula verdadera bajo una determinada condición, sino también cómo hacer falsa la misma fórmula bajo la misma condición. Por otra parte, fijando las fórmulas χ y ψ de L_{Lh}, la modalidad $[\chi, \psi]$ es una modalidad normal:

Proposición 8.17. *Sean χ, ψ, θ y φ fórmulas arbitrarias de L_{Lh}, la modalidad $[\chi, \psi]$ es normal:*

1. $\models [\chi, \psi](\theta \to \varphi) \to ([\chi, \psi]\theta \to [\chi, \psi]\varphi)$.
2. *Si* $\models \varphi$*, entonces* $\models [\chi, \psi]\varphi$.

Demostración. Tomando la definición del operador de aprender cómo, se tienen las siguientes igualdades:

$$[\alpha_1, \alpha_2]_i \alpha_3 = \neg \langle \alpha_1, \alpha_2 \rangle_i \neg \alpha_3 = \neg \langle \not\sim \rangle (\mathsf{Kh}_i(\alpha_1, \alpha_2) \wedge \neg \alpha_3) = [\not\sim](\mathsf{Kh}_i(\alpha_1, \alpha_2) \to \alpha_3).$$

Para el Ítem 1, sea $\mathfrak{M}$ un LTS^U. Supongamos que $\mathfrak{M}, w \models [\chi, \psi](\theta \to \varphi)$ y $\mathfrak{M}, w \models [\chi, \psi]\theta$. Luego $\mathfrak{M}, w \models [\not\sim](\mathsf{Kh}_i(\chi, \psi) \to (\theta \to \varphi))$ y $\mathfrak{M}, w \models [\not\sim](\mathsf{Kh}_i(\chi, \psi) \to \theta)$. Sean $\sigma_1, \sigma_2 \in \mathsf{Act}^*$, $\mathrm{U}'(i) \subseteq 2^{\mathsf{Act}^*}$ arbitrarios tales que $\mathrm{U}(i) \rightsquigarrow^{\sigma_1}_{\sigma_2} \mathrm{U}'(i)$. Si $\mathfrak{M}^{\mathrm{U}}_{\mathrm{U}'}, w \models \mathsf{Kh}_i(\chi, \psi)$, entonces $\mathfrak{M}^{\mathrm{U}}_{\mathrm{U}'}, w \models ((\theta \to \varphi) \wedge \theta)$. Por lo tanto, $\mathfrak{M}^{\mathrm{U}}_{\mathrm{U}'}, w \models \varphi$ y $\mathfrak{M}^{\mathrm{U}}_{\mathrm{U}'}, w \models (\mathsf{Kh}_i(\chi, \psi) \to \varphi)$. Con esto, $\mathfrak{M}, w \models [\not\sim](\mathsf{Kh}_i(\chi, \psi) \to \varphi)) = [\chi, \psi]\varphi$

Para el Ítem 2, $\models \varphi$ implica que $\models (\theta \to \varphi)$ para cualquier fórmula θ. Particularmente $\models (\mathsf{Kh}_i(\chi, \psi) \to \varphi)$. Usando la Proposición 8.10, $\models [\not\sim](\mathsf{Kh}_i(\chi, \psi) \to \varphi) = [\chi, \psi]\varphi$. □

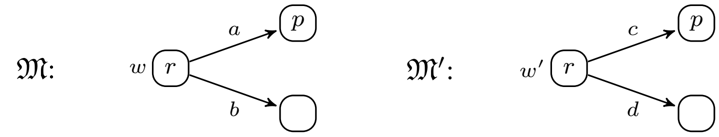
Por último, podemos establecer otros resultados de expresividad entre las modalidades dinámicas.

Proposición 8.18. *Las siguientes propiedades se cumplen:*

1. $\mathsf{L}_{\mathsf{Kh}_i} \prec \mathsf{L}_{\mathsf{Lh}}$.

2. $\mathsf{L}_{\mathsf{Ref}} \not\prec \mathsf{L}_{\mathsf{ARef}}$ *y* $\mathsf{L}_{\mathsf{Ref}} \not\prec \mathsf{L}_{\mathsf{Lh}}$.

Demostración. Considerando la Proposición 8.15, el Ítem 1 se demuestra con el modelo de las Proposiciones 8.8 y 8.13. En el mismo, la fórmula $\mathsf{L}_i(p,q)$ distingue a ambos LTS^{U}s. Para el Ítem 2 se tienen los siguientes LTS^{U}s con $\mathrm{U}(i) = \{\{a,b\}\}$ y $\mathrm{U}'(i) = \{\{c,d\}\}$:



Como L_{Lh} y $\mathsf{L}_{\mathsf{ARef}}$ no pueden hablar explícitamente sobre planes (más aún, por la Proposición 8.15, son equivalentes), $\mathfrak{M}, w$ y $\mathfrak{M}', w'$ son indistinguibles en ambas lógicas. En cambio, para $\mathsf{L}_{\mathsf{Ref}}$, $\mathfrak{M}, w \models \langle a \not\sim b \rangle \mathsf{Kh}_i(r,p)$ y $\mathfrak{M}', w' \not\models \langle a \not\sim b \rangle \mathsf{Kh}_i(r,p)$. $\square$

Concluyendo esta subsección, y por lo tanto este capítulo, en el mismo se han relevado diferentes extensiones dinámicas de $\mathsf{L}_{\mathsf{Kh}_i}$, ocupando operadores ónticos, que modifican la parte LTS de los modelos, y epistémicos, que modifican la incertidumbre de los agentes en U. Si bien para los primeros se han conseguido axiomas de reducción e incluso resultados de complejidad, nos hemos tenido que restringir a una clase muy específica de modelos. Mientras tanto, para los segundos, no se han podido encontrar clases de modelos que sean realmente significativas. Como mucho, se han considerado los LTS^{U}s tales que los agentes carecen de incertidumbre. Es decir, modelos tales que cada agente es capaz de distinguir cada plan que concibe como posible. Más aún, se tiene que sustitución uniforme no se cumple en general para estas lógicas dinámicas epistémicas. Considerando esta situación, en el siguiente capítulo se intentará compensar la gran desventaja de $\mathsf{L}_{\mathsf{Kh}_i}$, que es la de no poder hablar explícitamente de los planes testigo, introduciendo una modalidad que capaz de describir los planes de Act^*. Junto con este enriquecimiento del lenguaje base, se considerarán operadores dinámicos epistémicos que aprovecharán mejor esta nueva lógica.

Capítulo 9

Una extensión del lenguaje

En cada cosa es preciso preguntarse a sí mismo si es necesaria.

Marco Aurelio

Un problema recurrente a la hora de definir operadores dinámicos ónticos y epistémicos sobre la lógica $\mathsf{L}_{\mathsf{Kh}_i}$, es que resulta difícil obtener resultados de completitud y decidibilidad sobre la clase general de todos los LTS^{U}s. Esto se debe a que la expresividad de la lógica base no es suficiente para describir el efecto producido por estas modalidades y, por consiguiente, no resulta ser fácil encontrar axiomas de reducción. Esto tiene sentido, ya que la modalidad Kh_i es muy simple y no puede describir cursos de acción de manera explícita. Por esta razón, en el capítulo anterior se definieron algunas operaciones dinámicas de manera tal que se pudiera lograr completitud al menos sobre una clase restringida de modelos. Más aún, parece ser complejo conseguir completitud usando los métodos más convencionales, considerando ciertas características de estas lógicas, como lo es la falla de la regla de sustitución uniforme.

Por ello, en este capítulo se explorará una nueva alternativa. Extenderemos la lógica $\mathsf{L}_{\mathsf{Kh}_i}$ con una modalidad que puede hablar explícitamente de los cursos de acción que los agentes pueden tomar. Dicha modalidad es simplemente el operador de la lógica multimodal básica $[a]$ (con $a \in \mathsf{Act}$), que permite describir las acciones que son ejecutables en un estado. En primer lugar, extenderemos el sistema axiomático de la Tabla 5.1 para obtener una axiomatización completa para la nueva lógica. Además, con una definición de filtraciones similar a la usada en [24], se demuestra que esta lógica es decidible. Por último, introduciremos una serie de operadores dinámicos epistémicos. Los mismos reflejarán el hecho de que un agente se vuelve consciente de la existencia de un plan y elimina toda incertidumbre sobre el mismo. Las ventajas de este nuevo enfoque es que, dada la expresividad agregada en la lógica base, se podrán dar axiomas de reducción para estas modalidades dinámicas que valgan sobre la clase de todos los modelos. Este capítulo sirve como una extensión de lo visto en [9].

9.1. Sintaxis y semántica

En esta sección se introduce la sintaxis y la semántica de la lógica, denotada como $\mathsf{L}_{\mathsf{Kh}_i,\Box}$. La misma es una extensión de $\mathsf{L}_{\mathsf{Kh}_i}$ con la modalidad estándar $[a]$, parametrizada por una acción a perteneciente a Act [24, 26].

Definición 9.1. Las fórmulas de la lógica $\mathsf{L}_{\mathsf{Kh}_i,\Box}$ son las dadas por la siguiente gramática:

$$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid \mathsf{Kh}_i(\varphi, \varphi) \mid [a]\varphi,$$

con $p \in \mathsf{Prop}$, $i \in \mathsf{Agt}$ y $a \in \mathsf{Act}$. Las constantes y demás conectores booleanos $\bot$, $\top$, $\wedge$, $\rightarrow$, y $\leftrightarrow$ se definen de la forma usual. Se define la modalidad dual: $\langle a\rangle\varphi := \neg[a]\neg\varphi$. Las fórmulas de la forma $[a]\varphi$ se leen como *"cada ejecución de la acción a lleva siempre a situaciones en las que φ se cumple"*; mientras que $\langle a\rangle\varphi$ se lee como *"existe una situación después de ejecutar la acción a en la cual φ se cumple"*.

Se introduce la semántica usual de la modalidad $[a]$.

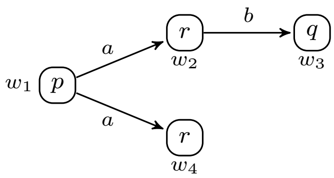
Bloque $\mathcal{L}$:	TAUT	$\vdash \varphi$ para φ una tautología proposicional
	DISTA	$\vdash \mathsf{A}(\varphi \to \psi) \to (\mathsf{A}\varphi \to \mathsf{A}\psi)$
	TA	$\vdash \mathsf{A}\varphi \to \varphi$
	4KhA	$\vdash \mathsf{Kh}_i(\psi, \varphi) \to \mathsf{A}\mathsf{Kh}_i(\psi, \varphi)$
	5KhA	$\vdash \neg\mathsf{Kh}_i(\psi, \varphi) \to \mathsf{A}\neg\mathsf{Kh}_i(\psi, \varphi)$
Bloque $\mathcal{L}_{\mathrm{LTS}^U}$:	KhA	$\vdash (\mathsf{A}(\chi \to \psi) \wedge \mathsf{Kh}_i(\psi, \varphi) \wedge \mathsf{A}(\varphi \to \theta)) \to \mathsf{Kh}_i(\chi, \theta)$
Bloque $\mathcal{L}_{\Box}$:	DIST$\Box$	$\vdash [a](\varphi \to \psi) \to ([a]\varphi \to [a]\psi)$
	A$\Box$	$\vdash \mathsf{A}\varphi \to [a]\varphi$
Reglas:	$\dfrac{\varphi \quad (\varphi \to \psi)}{\psi}$ MP	$\dfrac{\varphi}{\mathsf{A}\varphi}$ NECA

Tabla 9.1: Axiomatización $\mathcal{L}_{\mathsf{Kh}_i,\Box}$ de $\mathsf{L}_{\mathsf{Kh}_i,\Box}$ con respecto a los $\mathrm{LTS}^U s$.

Definición 9.2. Sea $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \mathrm{U}, \mathrm{V}\rangle$ un LTS^U, $a \in \mathsf{Act}$ y φ una fórmula en $\mathsf{L}_{\mathsf{Kh}_i,\Box}$, la semántica de $[a]$ y su dual $\langle a\rangle$ se definen de la siguiente forma:

$$\begin{array}{lll} \mathfrak{M}, w \models [a]\varphi & \text{sii} & \mathfrak{M}, v \models \varphi \text{ para todo } v \in \mathrm{R}_a(w), \\ \mathfrak{M}, w \models \langle a\rangle\varphi & \text{sii} & \mathfrak{M}, v \models \varphi \text{ para algún } v \in \mathrm{R}_a(w). \end{array}$$

Ejemplo 9.1. Sea $\mathfrak{M}$ el LTS^U del Ejemplo 4.2 con $\mathrm{U}(i) = \{\{b\}, \{ab, a\}\}$,



En este modelo se tiene que $\mathfrak{M}, w_1 \not\models \mathsf{Kh}(p, q)$ y $\mathfrak{M}, w_1 \not\models \mathsf{Kh}(p, r)$. Esto se debe a que no existe $\pi \in \mathrm{U}(i)$ tal que $[\![p]\!]^{\mathfrak{M}} = \{w_1\} \subseteq \mathrm{SE}(\pi)$ ya que $\mathrm{SE}(\{ab, a\}) = \mathrm{SE}(ab) \cap \mathrm{SE}(a) = \emptyset \cap \{w_1\} = \emptyset$ y $\mathrm{SE}(\{b\}) = \{w_2\}$. Convenientemente, podemos usar las nuevas modalidades para describir, por ejemplo, qué planes son fuertemente ejecutables en w_1, el único estado donde vale p. En el caso de $\{ab, a\}$, se cumple que $\mathfrak{M}, w_1 \models \langle a\rangle\top$ ($w_1 \in \mathrm{SE}(a)$) pero $\mathfrak{M}, w_1 \not\models [a]\langle b\rangle\top$ ($w_1 \notin \mathrm{SE}(ab)$). Mientras tanto, para $\{b\}$, se tiene que $\mathfrak{M}, w_1 \not\models \langle b\rangle\top$ ($w \notin \mathrm{SE}(b)$). $\Box$

Considerando el ejemplo anterior, dado un plan $\sigma \in \mathsf{Act}^*$, la descripción de si es fuertemente ejecutable o no en un determinado estado puede ser potencialmente expresada en el lenguaje $\mathsf{L}_{\mathsf{Kh}_i,\Box}$. Esto será muy últil a la hora de definir los operadores dinámicos de anuncios en las siguientes secciones, y obtener axiomatizaciones.

9.2. Sistema axiomático

La Tabla 9.1 introduce un sistema axiomático para $\mathsf{L}_{\mathsf{Kh}_i,\Box}$. El mismo es una extensión de la Tabla 5.1 con el agregado de dos axiomas del bloque $\mathcal{L}_{\Box}$: ***(1)*** DIST$\Box$, que establece que cada $[a]$ distribuye con respecto a la implicación, de la misma forma que lo hace la modalidad universal A en DISTA; y ***(2)*** A$\Box$, que describe que toda propiedad que se cumple globalmente también lo hace en los estados que estén vinculados por la acción a. Notar que la regla de NEC$\Box$, que aparece usualmente en la lógica modal, no es necesaria ya que se deriva usando NECA, A$\Box$ y MP. Es directo que estos axiomas y reglas son válidos en modelos arbitrarios y con esto la correctitud del sistema.

Antes de demostrar completitud, se deberán introducir algunas definiciones preliminares para el modelo canónico en el contexto de esta lógica, semejante a lo hecho en el Capítulo 5. Similarmente como en la Definición 5.1, se definen los conceptos de deducción, consistencia y maximalidad.

Definición 9.3. Sea $\Gamma \cup \{\varphi\}$ un conjunto fórmulas de $\mathsf{L}_{\mathsf{Kh}_i,\Box}$. Las nociones de deducción, $\Gamma \vdash \varphi$, y teorema, $\vdash \varphi$, se definen de la manera usual en el contexto de $\mathcal{L}_{\mathsf{Kh}_i,\Box}$. Un conjunto Γ es consistente con $\mathcal{L}_{\mathsf{Kh}_i,\Box}$ si $\Gamma \not\vdash \bot$, y Γ es maximal consistente con $\mathcal{L}_{\mathsf{Kh}_i,\Box}$ si Γ es consistente con $\mathcal{L}_{\mathsf{Kh}_i,\Box}$ y para todo $\varphi \notin \Gamma$, $\Gamma \cup \{\varphi\} \vdash \bot$.

Análogamente a Proposición 5.1, las propiedades KhE, SCOND y COND son derivables.

Proposición 9.1. *Las fórmulas* $(\mathsf{E}\psi \wedge \mathsf{Kh}_i(\psi,\varphi)) \to \mathsf{E}\varphi$ *(KhE),* $\mathsf{A}\neg\psi \to \mathsf{Kh}_i(\psi,\varphi)$ *(SCOND) y* $\mathsf{Kh}_i(\bot,\varphi)$ *(COND) son derivables de* $\mathcal{L}_{\mathsf{Kh}_i,\Box}$.
Es decir, (1) $\vdash (\mathsf{E}\psi \wedge \mathsf{Kh}_i(\psi,\varphi)) \to \mathsf{E}\varphi$ *(2)* $\vdash \mathsf{A}\neg\psi \to \mathsf{Kh}_i(\psi,\varphi)$ *y (3)* $\vdash \mathsf{Kh}_i(\bot,\varphi)$.

Definición 9.4. Sea $\mathbf{\Phi}$ el conjunto de todos los conjuntos maximales consistentes con $\mathcal{L}_{\mathsf{Kh}_i,\Box}$ de fórmulas en $\mathsf{L}_{\mathsf{Kh}_i,\Box}$. Para todo $\Delta \in \mathbf{\Phi}$, se definen:

$$\begin{aligned} \Delta|_{\mathsf{A}} &:= \{\mathsf{A}\psi \mid \mathsf{A}\psi \in \Delta\}, & \Delta|_{\neg\mathsf{A}} &:= \{\neg\mathsf{A}\psi \mid \neg\mathsf{A}\psi \in \Delta\},\\ \Delta|_{\mathsf{Kh}_i} &:= \{\mathsf{Kh}_i(\psi,\varphi) \mid \mathsf{Kh}_i(\psi,\varphi) \in \Delta\}, & \Delta|_{\neg\mathsf{Kh}_i} &:= \{\neg\mathsf{Kh}_i(\psi,\varphi) \mid \neg\mathsf{Kh}_i(\psi,\varphi) \in \Delta\},\\ \Delta|_{\mathsf{Kh}} &:= \textstyle\bigcup_{i\in\mathsf{Agt}} \Delta|_{\mathsf{Kh}_i}, & \Delta|_{\neg\mathsf{Kh}} &:= \textstyle\bigcup_{i\in\mathsf{Agt}} \Delta|_{\neg\mathsf{Kh}_i}. \end{aligned}$$

Sea Γ un conjunto en $\mathbf{\Phi}$. Se define $\mathsf{Act}_i^\Gamma := \{\langle\psi,\varphi\rangle \mid \mathsf{Kh}_i(\psi,\varphi) \in \Gamma\}$ para todo agente $i \in \mathsf{Agt}$, y $\mathsf{Act}^\Gamma := \bigcup_{i\in\mathsf{Agt}} \mathsf{Act}_i^\Gamma \cup \mathsf{Act}$.

Notar que como $\vdash \top$, se tiene que $\mathsf{A}\top = \mathsf{Kh}_i(\neg\top,\bot) \in \Gamma$, para todo $i \in \mathsf{Agt}$. Por lo tanto, $\mathsf{Act}^\Gamma \neq \emptyset$. Más aún, es numerable dado que Agt es finito y no vacío, y Act es numerable por definición. Con esta propiedad, usando una variación ligera de lo hecho en Capítulo 5, a continuación se considera la definición del modelo canónico.

Definición 9.5. Sea $\Gamma \in \mathbf{\Phi}$. La estructura $\mathfrak{M}^\Gamma = \langle \mathrm{S}^\Gamma, \mathrm{R}^\Gamma, \{\mathrm{U}^\Gamma(i)\}_{i\in\mathsf{Agt}}, \mathrm{V}^\Gamma\rangle$ sobre $\mathsf{Act}^\Gamma \cup \mathsf{Act}$, Agt y Prop se define de la siguiente manera.

- $\mathrm{S}^\Gamma := \{\Delta \in \mathbf{\Phi} \mid \Delta|_{\mathsf{A}} = \Gamma|_{\mathsf{A}}\}$;
- $\mathrm{R}^\Gamma_{\langle\psi,\varphi\rangle} := \bigcup_{i\in\mathsf{Agt}} \mathrm{R}^\Gamma_{\langle\psi,\varphi\rangle^i}$, con
$$\mathrm{R}^\Gamma_{\langle\psi,\varphi\rangle^i} = \{(\Delta_1,\Delta_2) \in (\mathrm{S}^\Gamma)^2 \mid \mathsf{Kh}_i(\psi,\varphi) \in \Gamma, \psi \in \Delta_1, \varphi \in \Delta_2\};$$
- $\mathrm{R}^\Gamma_a := \{(\Delta_1,\Delta_2) \in (\mathrm{S}^\Gamma)^2 \mid \text{para todo } [a]\varphi \in \Delta_1 \text{ implica } \varphi \in \Delta_2\}$;
- $\mathrm{U}^\Gamma(i) := \Big\{\{\langle\psi,\varphi\rangle\} \mid \langle\psi,\varphi\rangle \in \mathsf{Act}_i^\Gamma\Big\}$; y
- $\mathrm{V}^\Gamma(\Delta) := \{p \in \mathsf{Prop} \mid p \in \Delta\}$.

Notar que el modelo $\mathfrak{M}^\Gamma$ está generado por fórmulas de la forma $\mathsf{A}\varphi$, es decir, las fórmulas globales de Γ. Esto se debe a que ahora, a diferencia de lo ocurrido con $\mathsf{L}_{\mathsf{Kh}_i}$, contamos con más de una modalidad que genera nuestro modelo. Análogamente, se demuestra que $\mathfrak{M}^\Gamma$ es un LTS^U.

Proposición 9.2. *La estructura* $\mathfrak{M}^\Gamma = \langle \mathrm{S}^\Gamma, \mathrm{R}^\Gamma, \{\mathrm{U}^\Gamma(i)\}_{i\in\mathsf{Agt}}, \mathrm{V}^\Gamma\rangle$ *es un* LTS^U.

Demostración. Los conjuntos de acciones Act_i^Γ y Act son numerables. Por lo tanto su unión es numerable. Como $\Gamma \in \mathrm{S}^\Gamma$, $\mathrm{S}^\Gamma \neq \emptyset$. Es suficiente con mostrar que para cada $\mathrm{U}^\Gamma(i)$ se define una partición sobre un subconjunto no vacío de $2^{((\mathsf{Act}^\Gamma)^*)}$. Como $\mathsf{Kh}_i(\bot,\bot) \in \Gamma$, entonces $\langle\bot,\bot\rangle \in \mathsf{Act}_i^\Gamma$ y $\{\langle\bot,\bot\rangle\} \in \mathrm{U}^\Gamma(i)$; por lo tanto, $\mathrm{U}(i) \neq \emptyset$. Más aún, $\mathrm{U}(i)$ define una partición sobre $\bigcup_{\pi\in\mathrm{U}(i)} \pi$ ya que sus elementos conjuntos unitarios. Es directo que $\emptyset \notin \mathrm{U}^\Gamma(i)$. $\Box$

Notar que esta definición se aprovecha de la situación de que la modalidad Kh_i no puede expresar explícitamente el plan que sirve como testigo. Como corolario del "truth lemma" (Lema 9.1), dada una fórmula, se puede construir un modelo donde las acciones que el agente considera no aparecen en la fórmula.

Sea $\Gamma \in \mathbf{\Phi}$, las siguientes propiedades del modelo canónico $\mathfrak{M}^\Gamma$ serán importantes. Las pruebas son similares a las hechas en la Definición 5.1. Las acciones $\langle\psi,\varphi\rangle$ se referirán a los elementos en Act^Γ.

Proposición 9.3. *Para todo* $\Delta_1,\Delta_2 \in \mathrm{S}^\Gamma$ *se tiene que* $\Delta_1|_{\mathsf{X}} = \Delta_2|_{\mathsf{X}}$, $\Delta_1|_{\neg\mathsf{X}} = \Delta_2|_{\neg\mathsf{X}}$ *para* $\mathsf{X} \in \{\mathsf{Kh}_i, \mathsf{Kh}, \mathsf{A}\}$.

Demostración. $\Delta_1|_{\mathsf{A}} = \Delta_2|_{\mathsf{A}}$ es directo de la definición de S^Γ. Usando $\mathsf{4KhA}$ y TA, se tiene que $\vdash \mathsf{Kh}_i(\psi,\varphi) \to \mathsf{AKh}_i(\psi,\varphi)$. Como se está trabajando con conjuntos maximales consistentes y usando MP, $\mathsf{Kh}_i(\psi,\varphi) \in \Delta_1|_{\mathsf{Kh}_i}$ sii $\mathsf{AKh}_i(\psi,\varphi) \in \Delta_1|_{\mathsf{A}} = \Delta_2|_{\mathsf{A}}$. Esto equivale a que $\mathsf{Kh}_i(\psi,\varphi) \in \Delta_2|_{\mathsf{Kh}_i}$ y por lo tanto, $\Delta_1|_{\mathsf{Kh}_i} = \Delta_2|_{\mathsf{Kh}_i}$. Como consecuencia, $\Delta_1|_{\mathsf{Kh}} = \Delta_2|_{\mathsf{Kh}}$. Para $\Delta_1|_{\neg\mathsf{X}} = \Delta_2|_{\neg\mathsf{X}}$, las siguientes equivalencias se cumplen: $\neg\chi \in \Delta_1|_{\neg\mathsf{X}}$ sii $\chi \notin \Delta_1|_{\mathsf{X}}$ sii $\chi \notin \Delta_2|_{\mathsf{X}}$ sii $\neg\chi \in \Delta_2|_{\neg\mathsf{X}}$. □

Proposición 9.4. *Sea $\Delta \in \mathrm{S}^\Gamma$. Si Δ tiene un sucesor vía $\mathrm{R}^\Gamma_{\langle\psi,\varphi\rangle}$, entonces cada $\Delta' \in \mathrm{S}^\Gamma$ con $\varphi \in \Delta'$ puede ser alcanzado con $\mathrm{R}^\Gamma_{\langle\psi,\varphi\rangle}$ desde Δ.*

Demostración. Análoga a la Proposición 5.4. □

Proposición 9.5. *Sea φ una fórmula de $\mathsf{L}_{\mathsf{Kh}_i,\Box}$. Si $\varphi \in \Delta$ para todo $\Delta \in \mathrm{S}^\Gamma$, entonces $\mathsf{A}\varphi \in \Delta$ para todo $\Delta \in \mathrm{S}^\Gamma$.*

Demostración. Sea Δ en $\mathrm{S}^\Gamma \subseteq \mathbf{\Phi}$. Por definición, $(\Delta|_{\mathsf{A}} \cup \Delta|_{\neg\mathsf{A}}) \subseteq \Delta$, y por lo tanto es consistente. Más aún, para cualquier extensión maximalmente consistente Δ', debe cumplir que $\Delta|_{\mathsf{A}} = \Delta'|_{\mathsf{A}}$. Como es una extensión, $\Delta|_{\mathsf{A}} \subseteq \Delta'|_{\mathsf{A}}$. Si $\mathsf{A}\chi \notin \Delta|_{\mathsf{A}}$, entonces $\mathsf{A}\chi \notin \Delta$. Por lo tanto, $\neg\mathsf{A}\chi \in \Delta|_{\neg\mathsf{A}} \subseteq \Delta'|_{\neg\mathsf{A}} \subseteq \Delta'$ y con esto $\mathsf{A}\chi \notin \Delta'|_{\mathsf{A}}$.

Para demostrar la propiedad, supongamos que $\varphi \in \Delta$ para todo $\Delta \in \mathrm{S}^\Gamma$. Sea $\Delta \in \mathrm{S}^\Gamma$, $\Delta|_{\mathsf{A}} = \Gamma|_{\mathsf{A}}$. Luego el conjunto $\Delta|_{\mathsf{A}} \cup \Delta|_{\neg\mathsf{A}} \cup \{\neg\varphi\}$ es inconsistente. En caso contrario se podría extender a un conjunto maximal consistente $\Delta' \in \mathbf{\Phi}$. Por el resultado anterior, esto implica que $\Delta'|_{\mathsf{A}} = \Delta|_{\mathsf{A}}$, con lo cual $\Delta'|_{\mathsf{A}} = \Gamma|_{\mathsf{A}}$ y por lo tanto $\Delta' \in \mathrm{S}^\Gamma$. Por hipótesis, $\varphi \in \Delta'$, y por construcción, $\neg\varphi \in \Delta'$. Por lo tanto Δ' es inconsistente, una contradicción.

Con esto existen conjuntos $\{\mathsf{A}\psi_1,\ldots,\mathsf{A}\psi_n\} \subseteq \Delta|_{\mathsf{A}}$ y $\{\neg\mathsf{A}\psi'_1,\ldots,\neg\mathsf{A}\psi'_m\} \subseteq \Delta|_{\neg\mathsf{A}}$ tales que

$$\vdash \left(\bigwedge_{k=1}^{n} \mathsf{A}\psi_k \wedge \bigwedge_{k=1}^{m} \neg\mathsf{A}\psi'_k\right) \to \varphi.$$

Luego, por NECA,

$$\vdash \mathsf{A}\left(\left(\bigwedge_{k=1}^{n} \mathsf{A}\psi_k \wedge \bigwedge_{k=1}^{m} \neg\mathsf{A}\psi'_k\right) \to \varphi\right)$$

y usando DISTA y MP,

$$\vdash \mathsf{A}\left(\bigwedge_{k=1}^{n} \mathsf{A}\psi_k \wedge \bigwedge_{k=1}^{m} \neg\mathsf{A}\psi'_k\right) \to \mathsf{A}\varphi.$$

Ahora bien, usando $\mathsf{4KhA}$ y MP, $\mathsf{A}\psi_k \in \Delta|_{\mathsf{A}}$ implica que $\mathsf{AA}\psi_k \in \Delta$ para cada $k \in [1 \ldots n]$. Similarmente, usando $\mathsf{5KhA}$ y MP $\neg\mathsf{A}\psi'_k \in \Delta|_{\neg\mathsf{A}}$ implica que $\mathsf{A}\neg\mathsf{A}\psi'_k \in \Delta$ para cada $k \in [1 \ldots m]$. Por lo tanto,

$$\bigwedge_{k=1}^{n} \mathsf{AA}\psi_k \in \Delta \qquad \text{y} \qquad \bigwedge_{k=1}^{m} \mathsf{A}\neg\mathsf{A}\psi'_k \in \Delta,$$

y con esto

$$\bigwedge_{k=1}^{n} \mathsf{AA}\psi_k \wedge \bigwedge_{k=1}^{m} \mathsf{A}\neg\mathsf{A}\psi'_k \in \Delta, \quad \text{y se tiene que} \quad \mathsf{A}\left(\bigwedge_{k=1}^{n} \mathsf{A}\psi_k \wedge \bigwedge_{k=1}^{m} \neg\mathsf{A}\psi'_k\right) \in \Delta.$$

Por lo tanto, $\mathsf{A}\varphi \in \Delta$. □

Proposición 9.6. *Sean ψ, ψ', φ' fórmulas de $\mathsf{L}_{\mathsf{Kh}_i,\Box}$. Si para todo $\Delta \in \mathrm{S}^\Gamma$ tal que $\psi \in \Delta$, tiene un sucesor vía $\mathrm{R}^\Gamma_{\langle\psi',\varphi'\rangle}$. Entonces, $\mathsf{A}(\psi \to \psi') \in \Delta$ para todo $\Delta \in \mathrm{S}^\Gamma$.*

Demostración. Análoga a la Proposición 5.6 usando la Proposición 9.5. □

Proposición 9.7. *Sea $\Delta \in \mathrm{S}^\Gamma$ tal que $\{\psi, \mathsf{Kh}_i(\psi,\varphi)\} \subset \Delta$, entonces existe $\Delta' \in \mathrm{S}^\Gamma$ tal que $\varphi \in \Delta'$.*

Demostración. Análoga a la Proposición 5.7 usando las Proposiciones 9.5 y 9.1. □

Con estas propiedades se puede demostrar el "truth lemma" para $\mathfrak{M}^{\Gamma}$.

Lema 9.1. *Sean* $\Gamma \in \mathbf{\Phi}$*,* $\mathfrak{M}^{\Gamma} = \langle \mathrm{S}^{\Gamma}, \mathrm{R}^{\Gamma}, \{\mathrm{U}^{\Gamma}(i)\}_{i\in\mathsf{Agt}}, \mathrm{V}^{\Gamma}\rangle$*. Para todo* $\Theta \in \mathrm{S}^{\Gamma}$ *y* φ *fórmula de* $\mathsf{L}_{\mathsf{Kh}_i,\Box}$*,*

$$\mathfrak{M}^{\Gamma}, \Theta \models \varphi \qquad \textit{si y sólo si} \qquad \varphi \in \Theta.$$

Demostración. La prueba se hace por inducción en φ. Los casos atómicos y de los operadores booleanos son directos. El caso para Kh_i es análogo al visto en el Lema 5.1, por lo que sólo se demostrarán el caso para $[a]$.

- Caso $[a]\chi$: ($\Rightarrow$) Si $\mathfrak{M}^{\Gamma}, \Theta \models [a]\chi$, entonces por la definición de $\models$ y la hipótesis inductiva, para todo $\Delta \in \mathrm{R}^{\Gamma}_a(\Theta)$, $\chi \in \Delta$. Si $[a]\chi \notin \Theta$, entonces por consistencia maximal $\neg[a]\chi = \langle a\rangle\neg\chi \in \Theta$. Sea $\Delta^- := \Theta|_{\mathsf{A}} \cup \Theta|_{\neg\mathsf{A}} \cup \{\psi \mid [a]\psi \in \Theta\} \cup \{\neg\chi\}$. Tenemos que demostrar que Δ^- es consistente. Caso contrario, existen conjuntos $\{\mathsf{A}\psi_1, \ldots, \mathsf{A}\psi_n\} \subseteq \Delta|_{\mathsf{A}}$, $\{\neg\mathsf{A}\psi'_1, \ldots, \neg\mathsf{A}\psi'_m\} \subseteq \Delta|_{\neg\mathsf{A}}$ y $\{\psi''_1, \ldots, \psi''_l\} \subseteq \{\psi \mid [a]\psi \in \Theta\}$ tales que:

  $$\vdash \left(\bigwedge_{k=1}^{n} \mathsf{A}\psi_k \wedge \bigwedge_{k=1}^{m} \neg\mathsf{A}\psi'_k \wedge \bigwedge_{k=1}^{l} \psi''_k\right) \rightarrow \chi.$$

  Por NECA, $\mathsf{A}\Box$ y $\mathsf{DIST}\Box$ se tiene que:

  $$\vdash [a]\left(\bigwedge_{k=1}^{n} \mathsf{A}\psi_k \wedge \bigwedge_{k=1}^{m} \neg\mathsf{A}\psi'_k \wedge \bigwedge_{k=1}^{l} \psi''_k\right) \rightarrow [a]\chi.$$

  Dado que $\vdash [a](\varphi_1 \wedge \varphi_2) \rightarrow ([a]\varphi_1 \wedge [a]\varphi_2)$, esto es equivalente a:

  $$\vdash \left(\bigwedge_{k=1}^{n} [a]\mathsf{A}\psi_k \wedge \bigwedge_{k=1}^{m} [a]\neg\mathsf{A}\psi'_k \wedge \bigwedge_{k=1}^{l} [a]\psi''_k\right) \rightarrow [a]\chi.$$

  Por construcción, $[a]\psi''_k \in \Theta$. Dado que por hipótesis $\mathsf{A}\psi_k \in \Theta$, utilizando $\mathsf{4KhA}$, se tiene que $\mathsf{AA}\psi_k \in \Theta$. Aplicando $\mathsf{A}\Box$ ($\mathsf{AA}\psi_k \rightarrow [a]\mathsf{A}\psi_k$), se obtiene que $[a]\mathsf{A}\psi_k \in \Theta$. Con un argumento similar con $\mathsf{5KhA}$, $[a]\neg\mathsf{A}\psi'_k \in \Theta$. Por lo tanto, $[a]\chi \in \Theta$, una contradicción.

  Como Δ^- es consistente, entonces se tiene una expansión maximal consistente Δ tal que $\Delta|_{\mathsf{A}} = \Theta|_{\mathsf{A}}$ (por lo tanto, $\Delta \in \mathrm{S}^{\Gamma}$), $\Delta \in \mathrm{R}^{\Gamma}_a(\Theta)$ y $\chi \notin \Delta$. Pero eso es una contradicción. Luego, $[a]\chi \in \Theta$.

  ($\Leftarrow$) Supongamos que $[a]\chi \in \Theta$. Tomemos $\Delta \in \mathrm{R}^{\Gamma}_a(\Theta)$, por definición de R^{Γ}_a, $\chi \in \Delta$. Por hipótesis inductiva y la definición de $\models$, para todo $\Delta \in \mathrm{R}^{\Gamma}_a(\Theta)$, se tiene que $\mathfrak{M}^{\Gamma}, \Delta \models \chi$. Por lo tanto, $\mathfrak{M}^{\Gamma}, \Theta \models [a]\chi$.

□

Usando el lema anterior, se finaliza esta sección con la demostración de correctitud y completitud para esta axiomatización.

Teorema 9.1. *El sistema axiomático* $\mathcal{L}_{\mathsf{Kh}_i,\Box}$ *(Tabla 9.1) es correcto y fuertemente completo para* $\mathsf{L}_{\mathsf{Kh}_i,\Box}$ *con respecto a la clase de todos los* LTS^U*s.*

Demostración. Para la correctitud, basta con mostrar que los axiomas del sistema son validos y que sus reglas preservan la validez, lo cual resulta ser directo. Para la completitud fuerte, sea Γ' un conjunto de fórmulas consistente con $\mathcal{L}_{\mathsf{Kh}_i,\Box}$. Teniendo en cuenta la Observación 5.2 aplicada a $\mathsf{L}_{\mathsf{Kh}_i,\Box}$, Γ' puede ser extendido a un conjunto maximal consistente con $\mathcal{L}_{\mathsf{Kh}_i,\Box}$ $\Gamma \supseteq \Gamma'$. Por el Lema 9.1, $\mathfrak{M}^{\Gamma}, \Gamma \models \Gamma'$ y Γ' es satisfacible. Dado que $\mathfrak{M}^{\Gamma}$ es un LTS^U (Proposición 9.2), esto termina de demostrar la completitud fuerte. □

9.3. Decidibilidad vía filtraciones

De la misma manera que en la Sección 7.1 se demostró decidibilidad para $\mathsf{L}_{\mathsf{Kh}_i}$ utilizando filtraciones, demostraremos decidibilidad para $\mathsf{L}_{\mathsf{Kh}_i,\Box}$. Para ello, se vuelven a introducir las relaciones $\leftrightsquigarrow_\Sigma$ y $\leftrightarrows_\Sigma$ de la Definición 7.1. Notar que para este caso, no es necesario agregar nuevas condiciones para estas relaciones. Esto es debido a que $\leftrightsquigarrow_\Sigma$ sigue hablando de equivalencia de subfórmulas y $\leftrightarrows_\Sigma$ habla de los testigos de las fórmulas Kh_i. Más aún, sin la parte de planes, esta sería la definición estándar de filtraciones para una lógica multimodal [24, 26].

Definición 9.6 (Σ-equivalencia)**.** Sea $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \{\mathrm{U}(i)\}_{i\in\mathsf{Agt}}, \mathrm{V}\rangle$ un LTS^U y Σ un conjunto de fórmulas de $\mathsf{L}_{\mathsf{Kh}_i,\Box}$ cerrada bajo subfórmulas. Se definen las relaciones de equivalencia $\leftrightsquigarrow_\Sigma \subseteq \mathrm{S}\times\mathrm{S}$ y $\leftrightarrows_\Sigma \subseteq \mathrm{U}_{\mathsf{Agt}} \times \mathrm{U}_{\mathsf{Agt}}$ (con $\mathrm{U}_{\mathsf{Agt}} := \bigcup_{i\in\mathsf{Agt}} \mathrm{U}(i)$) como:

$$\begin{array}{ll} w \leftrightsquigarrow_\Sigma v & \text{sii para todo } \psi \in \Sigma,\ \mathfrak{M}, w \models \psi \text{ sii } \mathfrak{M}, v \models \psi, \\ \pi \leftrightarrows_\Sigma \pi' & \text{sii para todo } i \in \mathsf{Agt} \text{ y } \mathsf{Kh}_i(\psi,\varphi) \in \Sigma,\ \pi \text{ es un testigo para} \\ & \quad \mathsf{Kh}_i(\psi,\varphi) \text{ en } \mathfrak{M} \text{ sii } \pi' \text{ es un testigo para } \mathsf{Kh}_i(\psi,\varphi) \text{ en } \mathfrak{M}. \end{array}$$

Dado $w \in \mathrm{S}$ y $\pi \in \mathrm{U}_{\mathsf{Agt}}$, se definen las clases de equivalencia sobre Σ:

$$[w]_\Sigma := \{v \in \mathrm{S} \mid w \leftrightsquigarrow_\Sigma v\}; \qquad [\pi]_\Sigma := \{\pi' \in \mathrm{U}_{\mathsf{Agt}} \mid \pi \leftrightarrows_\Sigma \pi'\}.$$

Análogamente a la Definición 7.2, se establecen las acciones básicas que se utilizarán en la filtración. En este caso, se considerarán además todas acciones básicas de Act que ocurran en Σ.

Definición 9.7. Sea $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \{\mathrm{U}(i)\}_{i\in\mathsf{Agt}}, \mathrm{V}\rangle$ un LTS^U definido sobre un conjunto numerable de acciones Act, y sea Σ un conjunto de fórmulas de $\mathsf{L}_{\mathsf{Kh}_i,\Box}$ cerrado bajo subfórmulas. Se definen

- $\mathsf{Act}_i^\Sigma := \{a_{[\pi]_\Sigma} \mid \pi \in \mathrm{U}(i)$ es un testigo para algún $\mathsf{Kh}_i(\psi,\varphi) \in \Sigma$ en $\mathfrak{M}\}$,
- $\mathsf{Act}_\Box^\Sigma := \{a \in \mathsf{Act} \mid [a]\varphi \in \Sigma\}$, y
- $\mathsf{Act}^\Sigma := \bigcup_{i\in\mathsf{Agt}} \mathsf{Act}_i^\Sigma \cup \mathsf{Act}_\Box^\Sigma$.

Con esto, damos la definición de filtraciones vía un conjunto de fórmulas Σ cerrado bajo subfórmulas, similar a la dada en la Definición 7.3, considerando ahora los operadores $[a]$.

Definición 9.8 (Filtración de $\mathfrak{M}$ vía Σ)**.** Sean $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \{\mathrm{U}(i)\}_{i\in\mathsf{Agt}}, \mathrm{V}\rangle$ un LTS^U definido sobre Act y Σ un conjunto de fórmulas de $\mathsf{L}_{\mathsf{Kh}_i,\Box}$ cerrado bajo subfórmulas. Un LTS^U $\mathfrak{M}^f = \langle \mathrm{S}^f, \mathrm{R}^f, \{\mathrm{U}^f(i)\}_{i\in\mathsf{Agt}}, \mathrm{V}^f\rangle$ definido sobre $(\mathsf{Act}^\Sigma \cup \{a_\emptyset\})$ es una *filtración de $\mathfrak{M}$ vía Σ* si y sólo si se cumplen las siguientes condiciones:

1. $\mathrm{S}^f := \{[w]_\Sigma \mid w \in \mathrm{S}\}$;
2. $\mathrm{V}^f([w]_\Sigma) := \{p \in \Sigma \mid \mathfrak{M}, w \models p\}$;
3. $\mathrm{U}(i)^f := \{\{a_{[\pi]_\Sigma}\} \mid a_{[\pi]_\Sigma} \in \mathsf{Act}_i^\Sigma\} \cup \{\{a_\emptyset\}\}$ para cada $i \in \mathsf{Agt}$;
4. $\mathrm{R}^f_{a_\emptyset} = \emptyset$ y para todo $a_{[\pi]_\Sigma} \in \mathsf{Act}_i^\Sigma$, existe un subconjunto no vacío $G_{[\pi]_\Sigma} \subseteq [\pi]_\Sigma$ tal que $([w]_\Sigma, [v]_\Sigma) \in \mathrm{R}^f_{a_{[\pi]_\Sigma}}$ sii
   - para todo $w' \in [w]_\Sigma$ y $\pi' \in G_{[\pi]_\Sigma}$, se tiene que $w' \in \mathrm{SE}(\pi')$, y
   - existen $w'' \in [w]_\Sigma$, $v'' \in [v]_\Sigma$ y $\pi'' \in G_{[\pi]_\Sigma}$ tales que $(w', v') \in \mathrm{R}_{\pi''}$;
5. para todo $a \in \mathsf{Act}_\Box^\Sigma$, si $(w, v) \in \mathrm{R}_a$, entonces $([w]_\Sigma, [v]_\Sigma) \in \mathrm{R}^f_a$; y
6. para todo $a \in \mathsf{Act}_\Box^\Sigma$, si $([w]_\Sigma, [v]_\Sigma) \in \mathrm{R}^f_a$ entonces para todo $[a]\varphi \in \Sigma$, si $\mathfrak{M}, w \models [a]\varphi$ entonces $\mathfrak{M}, v \models \varphi$.

Similarmente a lo que sucedió con la Definición 7.3, la Definición 9.8 requiere algunas aclaraciones. Para la parte LTS, la filtración se define de manera similar a la que se hace en la lógica modal básica para S^f y V^f [24]. $\mathrm{R}^f_{a_{[\pi]_\Sigma}}$ se define con respecto a un subconjunto $G_{[\pi]_\Sigma} \subseteq [\pi]_\Sigma$, cuyos elementos coinciden en ser testigos de las mismas fórmulas Kh_i en Σ (Ítem 4). Para una clase de equivalencia $[w]_\Sigma$, se añaden arcos sólo si todo $\pi' \in G_{[\pi]_\Sigma}$ es fuertemente ejecutable en cada estado $w' \in [w]_\Sigma$. $\mathrm{S}^f \neq \emptyset$ y V^f está bien definida: dado $p \in \Sigma$, si $[w]_\Sigma = [v]_\Sigma$ y $\mathfrak{M}, w \models p$, entonces $\mathfrak{M}, v \models p$. Se puede ver que $\mathrm{U}^f(i)$ también está bien definida por construcción (Observación 4.1). Especialmente, de no existir ningún $\pi \in \mathrm{U}(i)$ tal que cumpla algún Kh_i en Σ, se sigue garantizando que $\mathrm{U}^f(i) \neq \emptyset$. Esto se debe a que se considera una acción vacua, $a_\emptyset$, que no es fuertemente ejecutable en ningún estado. Además, considerando que $(\mathsf{Act}^\Sigma \cup \{a_\emptyset\})$ es numerable, $\mathfrak{M}^f$ está bien definida. Se tiene además que las acciones del agente i en $\mathrm{U}^f(i)$ están separadas de las acciones definidas en Act. Más aún, estas acciones nuevas $a_{[\pi]_\Sigma}$ que se introducen como resultado de que exista un testigo para algún Kh_i, no impactan en ninguna fórmula de Σ y no generan conflicto. Por otro lado, las condiciones de los Ítems 5 y 6 son las usuales en la lógica modal básica [24, 26]. Con esto, procedemos a demostrar que, dado un modelo y su correspondiente filtración vía Σ, ambos satisfacen las mismas fórmulas de Σ.

Teorema 9.2. *Sea $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \{\mathrm{U}(i)\}_{i \in \mathsf{Agt}}, \mathrm{V}\rangle$ un LTS^U y sea Σ un conjunto de fórmulas de $\mathsf{L}_{\mathsf{Kh}_i,\Box}$ cerrado bajo subfórmulas. Entonces, para todo $\psi \in \Sigma$ y $w \in \mathrm{S}$, $\mathfrak{M}, w \models \psi$ sii $\mathfrak{M}^f, [w]_\Sigma \models \psi$. Más aún, si Σ es finito entonces $\mathfrak{M}^f$ es un modelo finito.*

Demostración. Los casos base y booleanos son bastante directos, ya sea usando la definición de $\models$ o la hipótesis inductiva. Por otra parte, el caso para Kh_i se demostró en el Teorema 7.1. Se procederá con el caso para la modalidad $[a]$. Supongamos que $\mathfrak{M}, w \models [a]\varphi$. Sea $([w]_\Sigma, [v]_\Sigma) \in \mathrm{R}^f_a$, como $[a]\varphi \in \Sigma$, entonces $\mathfrak{M}, v \models \varphi$. Por hipótesis inductiva, $\mathfrak{M}^f, [v]_\Sigma \models \varphi$. Como se demostró para un $[v]_\Sigma \in \mathrm{S}^f$ arbitrario, se tiene que $\mathfrak{M}^f, [w]_\Sigma \models [a]\varphi$. Supongamos que $\mathfrak{M}^f, [w]_\Sigma \models [a]\varphi$. Sea $(w, v) \in \mathrm{R}_a$, entonces $([w]_\Sigma, [v]_\Sigma) \in \mathrm{R}^f_a$ y $\mathfrak{M}^f, [v]_\Sigma \models \varphi$. Por hipótesis inductiva, $\mathfrak{M}, v \models \varphi$. Como se demostró para un $v \in \mathrm{S}$ arbitrario, se tiene que $\mathfrak{M}, w \models [a]\varphi$.

Queda demostrar que si Σ es finito, entonces $\mathfrak{M}^f$ también lo es. En primer lugar, el número de estados en S^f es a lo sumo 2^m, siendo m el número total de fórmulas en Σ. Por lo tanto, S^f y V^f son finitos. Por definición, para todo $i \in \mathsf{Agt}$, Act^Σ_i es a lo sumo exponencial en el número de $\mathsf{Kh}_i(\psi, \varphi) \in \Sigma$. Como además $\mathsf{Act}^\Sigma_\Box$ es polinomial en el número de $[a]\varphi \in \Sigma$, $\mathsf{Act}^\Sigma \cup \{a_\emptyset\}$ es finito. Con esto, R^f y U^f son finitos, y $\mathfrak{M}^f$ es finito. □

Considerando el teorema anterior, basta entonces con demostrar que, dado un modelo $\mathfrak{M}$ y un conjunto de fórmulas Σ cerrado bajo subfórmulas, el conjunto de filtraciones de $\mathfrak{M}$ vía Σ es no vacío. Esto se puede ver definiendo $G_{[\pi]_\Sigma} = [\pi]_\Sigma \neq \emptyset$ para cada $a_{[\pi]_\Sigma} \in \mathsf{Act}^\Sigma_i$, y considerando al menos una de dos alternativas para definir R_a para cada $a \in \mathsf{Act}^\Sigma_\Box$ dadas en [24, 26]:

- $([w]_\Sigma, [v]_\Sigma) \in \mathrm{R}^f_a$ sii existen $w' \in [w]_\Sigma$, $v' \in [v]_\Sigma$ tales que $(w', v') \in \mathrm{R}_a$; ó
- $([w]_\Sigma, [v]_\Sigma) \in \mathrm{R}^f_a$ sii para todo $[a]\varphi \in \Sigma$, $\mathfrak{M}, w \models [a]\varphi$ entonces $\mathfrak{M}, v \models \varphi$.

Como consecuencia, el problema de satisfacibilidad para esta lógica es decidible.

Corolario 9.1. *El problema de satisfacibilidad para $\mathsf{L}_{\mathsf{Kh}_i,\Box}$ es decidible.*

Demostración. La demostración es similar a la vista en el Corolario 7.1. □

El resto de las secciones de este capítulo se enfocarán en introducir una familia de lógicas dinámicas epistémicas que extienden a $\mathsf{L}_{\mathsf{Kh}_i,\Box}$. A su vez, dichas extensiones cuentan con axiomas de reducción y, por lo tanto, axiomatizaciones correctas y fuertemente completas sobre la clase de todos los modelos. Más aún, considerando el Corolario 9.1, se tiene que el problema de satisfacibilidad para cada una de estas es decidible.

9.4. Refinamiento con acciones atómicas

El primer operador que consideraremos está inspirado en la modalidad de refinamiento de la Subsección 8.2.1. El mismo será denotado como $[!a]$, para $a \in \mathsf{Act}$, y lo interpretaremos como el anuncio público (es decir, para todos los agentes) de que la acción a puede ser distinguida de todos los demás posibles planes. A diferencia de la modalidad $[\sigma_1 \not\sim \sigma_2]$, esta modalidad toma acciones atómicas (esto será generalizado luego para planes arbitrarios), y tiene un único parámetro en vez de dos. Con este operador, se aísla el efecto de una acción, en vez de separar una clase de acuerdo al efecto de dos acciones. Esto resulta útil para, por ejemplo, que un agente pueda recibir información concreta sobre el efecto de una acción particular, haciendo que esta acción se vuelva distinguible de cualquier otra.

Definición 9.9. Las fórmulas de la lógica $\mathsf{L}_{\mathsf{Kh}_i,\Box,[!a]}$ son las dadas por la siguiente gramática:

$$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid \mathsf{Kh}_i(\varphi,\varphi) \mid [a]\varphi \mid [!a]\varphi,$$

con $p \in \mathsf{Prop}$, $i \in \mathsf{Agt}$ y $a \in \mathsf{Act}$. Las fórmulas de la forma $[!a]\varphi$ se leen como *"después de anunciar que a es distinguible de cualquier otro plan, φ se cumple".*

Definición 9.10. Sean $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \mathrm{U}, \mathrm{V}\rangle$ un LTS^U, $a \in \mathsf{Act}^*$ y φ una fórmula de $\mathsf{L}_{\mathsf{Kh}_i,\Box,[!\sigma]}$,

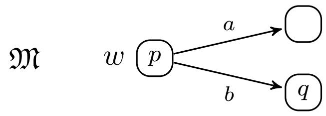
$$\mathfrak{M}, w \models [!a]\varphi \quad\Leftrightarrow\quad \mathfrak{M}^a, w \models \varphi,$$

donde $\mathfrak{M}^a = \langle \mathrm{S}, \mathrm{R}, \mathrm{U}', \mathrm{V}\rangle$, con:

- $\mathrm{U}'(i) = (\mathrm{U}(i) \setminus \{\pi\}) \cup \{\{a\}\}$ si existe $\pi \in \mathrm{U}(i)$ tal que $a \in \pi$,
- $\mathrm{U}'(i) = \mathrm{U}(i) \cup \{\{a\}\}$ en caso contrario.

Esta nueva operación dinámica es un anuncio de que una acción básica tiene un comportamiento que puede ser identificado de los demás planes, en algunos casos descartando planes de la percepción de los agentes. En caso de que la acción no sea parte de las consideradas para un determinado agente, se la agrega haciendo que el agente sea consciente de su existencia. Consideremos el siguiente ejemplo.

Ejemplo 9.2. Sea $\mathfrak{M}$ un modelo de un solo agente representado abajo con $\mathrm{U}(i) := \{\{a\}\}$.



Claramente, $\mathfrak{M}, w \not\models \mathsf{Kh}_i(p,q)$. Pero $\mathfrak{M}, w \models [!b]\mathsf{Kh}_i(p,q)$, dado que al anunciar b, el agente se vuelve consciente de la existencia del plan, y a raíz de ello sabe cómo garantizar q dado p. $\Box$

Antes de ver la axiomatización de esta extensión, se tiene la siguiente proposición establece que la modalidad es normal y su dual es la misma modalidad.

Proposición 9.8. *Las siguientes propiedades se cumplen:*

1. $\models [!a]\varphi \leftrightarrow \neg[!a]\neg\varphi$.
2. $\models [!a](\varphi_1 \to \varphi_2) \to ([!a]\varphi_1 \to [!a]\varphi_2)$.
3. *Si* $\models \varphi$, *entonces* $\models [!a]\varphi$.

Demostración. Análoga a la de las Proposiciones 8.2 y 8.4. $\Box$

Sistema axiomático

A diferencia de los operadores dinámicos epistémicos del Capítulo 8, en este caso es posible dar una axiomatización completa para la lógica $\mathsf{L}_{\mathsf{Kh}_i,\Box,[!a]}$. Si bien generalmente lidiar con modalidades dinámicas no es fácil, el poder expresivo de la lógica base ($\mathsf{L}_{\mathsf{Kh}_i,\Box}$) es suficiente para capturar el comportamiento de $[!a]$ y obtener una axiomatización vía axiomas de reducción (Tabla 9.2).

Análogamente a las Subsecciones 8.1.1 y 8.1.2, la demostración de completitud consta de usar axiomas de reducción que traducen cada fórmula de $\mathsf{L}_{\mathsf{Kh}_i,\Box,[!a]}$ en una fórmula equivalente de $\mathsf{L}_{\mathsf{Kh}_i,\Box}$. Esto se hace por inducción en la estructura de las fórmulas de $\mathsf{L}_{\mathsf{Kh}_i,\Box,[!a]}$. La dificultad, como siempre, es eliminar todas las ocurrencias de la modalidad $[!a]$ en presencia de un Kh_i. Consideremos la fórmula $[!a]\mathsf{Kh}_i(\psi,\varphi)$. Después del "anuncio" de que la acción a es diferente a cualquier otro plan, existen dos posibilidades de por qué $\mathsf{Kh}_i(\psi,\varphi)$ se cumple. La primera es que la acción a no cumple ningún rol en el valor de verdad de $\mathsf{Kh}_i(\psi,\varphi)$. En ese caso, se puede poner la modalidad dinámica en la pre y postcondición de la modalidad Kh_i y proceder recursivamente con su eliminación. Es decir, se obtiene $\mathsf{Kh}_i([!a]\psi,[!a]\varphi)$. La segunda posibilidad es que a es importante para generar nuevo conocimiento para el agente. Si el conjunto unitario $\{a\}$ es el testigo para $\mathsf{Kh}_i(\psi,\varphi)$ es porque para todo estado que satisfaga ψ después del anuncio

RAtom	$\vdash [!a]p \leftrightarrow p$
R¬	$\vdash [!a]\neg\varphi_1 \leftrightarrow \neg[!a]\varphi_1$
R∨	$\vdash [!a](\varphi_1 \vee \varphi_2) \leftrightarrow ([!a]\varphi_1 \vee [!a]\varphi_2)$
R□	$\vdash [!a][a]\varphi_1 \leftrightarrow [a][!a]\varphi_1$
RKh	$\vdash [!a]\mathsf{Kh}_i(\varphi_1, \varphi_2) \leftrightarrow (\mathsf{Kh}_i([!a]\varphi_1, [!a]\varphi_2) \vee \mathsf{A}([!a]\varphi_1 \rightarrow (\langle a\rangle\top \wedge [a][!a]\varphi_2)))$
$\mathsf{RE}_{[!]}$	De $\vdash \varphi \leftrightarrow \psi$ se deriva $\vdash [!a]\varphi \leftrightarrow [!a]\psi$

Tabla 9.2: Axiomas de reducción $\mathcal{L}_{[!a]}$.

de a ($[!a]\psi$): ***(1)*** a es fuertemente ejecutable en dicho estado, y ***(2)*** después de su ejecución se cumple la postcondición φ en el modelo actualizado por $[!a]$ ($[!a]\varphi$).

El primer ítem se lo captura mediante la fórmula $[!a]\psi \rightarrow \langle a\rangle\top$ que se debe cumplir en todos los estados (notar que ejecutabilidad fuerte es trivial si a puede ejecutarse, ya que se trata de una acción atómica). Mientras tanto, el segundo ítem se lo puede formalizar con $[a][!a]\varphi$. Juntando estas dos condiciones, este caso se lo describe con la fórmula $\mathsf{A}([!a]\psi \rightarrow (\langle a\rangle\top \wedge [a][!a]\varphi))$. Los demás casos son reducciones estándar ya vistas en las Tablas 8.1 y 8.2 para los operadores de anuncios públicos y actualización de relaciones respectivamente. Con esto, se obtienen los axiomas de la Tabla 9.2.

Lema 9.2. *Los axiomas de reducción de la Tabla 9.2 son válidos.*

Demostración. Dado que la semántica de $[!a]$ sólo modifica la perspectiva de los agentes (U), los axiomas de reducción RAtom, R¬, R∨ y R□ se cumplen usando la definición de $\models$ (la demostración es similar a la del Lema 8.2). Por otro lado, $\mathsf{RE}_{[!]}$ se cumple utilizando la normalidad de $[!a]$ demostrada en la Proposición 9.8. Por lo tanto, nos queda mostrar que RKh es válida. Sea $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \mathrm{U}, \mathrm{V}\rangle$ un LTS^U.

($\Rightarrow$): $\mathfrak{M} \models [!a]\mathsf{Kh}_i(\varphi_1, \varphi_2)$ sii $\mathfrak{M}^a \models \mathsf{Kh}_i(\varphi_1, \varphi_2)$ sii existe $\pi' \in \mathrm{U}'(i)$ tal que $[\![\varphi_1]\!]^{\mathfrak{M}^a} \subseteq \mathrm{SE}^{\mathfrak{M}^a}(\pi')$ y $\mathrm{R}_{\pi'}([\![\varphi_1]\!]^{\mathfrak{M}^a}) \subseteq [\![\varphi_2]\!]^{\mathfrak{M}^a}$. Usando que $[\![\psi]\!]^{\mathfrak{M}^a} = [\![[!a]\psi]\!]^{\mathfrak{M}}$ y $\mathrm{SE}^{\mathfrak{M}^a}(\pi') = \mathrm{SE}^{\mathfrak{M}}(\pi')$ (las relaciones no cambian), esto es equivalente a que exista $\pi' \in \mathrm{U}'(i)$ tal que $[\![[!a]\varphi_1]\!]^{\mathfrak{M}} \subseteq \mathrm{SE}^{\mathfrak{M}}(\pi')$ y $\mathrm{R}_{\pi'}([\![[!a]\varphi_1]\!]^{\mathfrak{M}}) \subseteq [\![[!a]\varphi_2]\!]^{\mathfrak{M}}$. Se consideran dos casos:

- Si $\pi' \neq \{a\}$, entonces $\pi' \in \mathrm{U}(i)$ y se tiene que $\mathfrak{M} \models \mathsf{Kh}_i([!a]\varphi_1, [!a]\varphi_2)$.
- Si $\pi' = \{a\}$, entonces $[\![[!a]\varphi_1]\!]^{\mathfrak{M}} \subseteq \mathrm{SE}^{\mathfrak{M}}(a)$ y $\mathrm{R}_a([\![[!a]\varphi_1]\!]^{\mathfrak{M}}) \subseteq [\![[!a]\varphi_2]\!]^{\mathfrak{M}}$. De la primera parte, usando la definición de ejecutabilidad fuerte, se deriva que para todo $w \in [\![[!a]\varphi_1]\!]^{\mathfrak{M}}$, se tiene $\mathfrak{M}, w \models \langle a\rangle\top$. De la segunda parte, se tiene que para todo $w \in [\![[!a]\varphi_1]\!]^{\mathfrak{M}}$, $\mathfrak{M}, w \models [a][!a]\varphi_2$. Por lo tanto, para todo $w \in \mathrm{S}$, $\mathfrak{M}, w \models [!a]\varphi_1 \rightarrow (\langle a\rangle\top \wedge [a][!a]\varphi_2)$.

($\Leftarrow$): $\mathfrak{M} \models \mathsf{Kh}_i([!a]\varphi_1, [!a]\varphi_2) \vee \mathsf{A}([!a]\varphi_1 \rightarrow (\langle a\rangle\top \wedge [a][!a]\varphi_2))$, entonces la disyunción se cumple para alguna de las dos situaciones:

- $\mathfrak{M} \models \mathsf{Kh}_i([!a]\varphi_1, [!a]\varphi_2)$: Sea $\pi \in \mathrm{U}(i)$ tal que $[\![[!a]\varphi_1]\!]^{\mathfrak{M}} \subseteq \mathrm{SE}^{\mathfrak{M}}(\pi)$ y $\mathrm{R}_\pi([\![[!a]\varphi_1]\!]^{\mathfrak{M}}) \subseteq [\![[!a]\varphi_2]\!]^{\mathfrak{M}}$. Como se mencionó anteriormente, esto es equivalente a $[\![\varphi_1]\!]^{\mathfrak{M}^a} \subseteq \mathrm{SE}^{\mathfrak{M}^a}(\pi)$ y $\mathrm{R}_\pi([\![\varphi_1]\!]^{\mathfrak{M}^a}) \subseteq [\![\varphi_2]\!]^{\mathfrak{M}^a}$. Si $a \notin \pi$, entonces $\pi \in \mathrm{U}'(i)$. Si $a \in \pi$, entonces $\{a\} \in \mathrm{U}'(i)$, y $\mathrm{SE}^{\mathfrak{M}^a}(\pi) \subseteq \mathrm{SE}^{\mathfrak{M}^a}(\{a\})$ y $\mathrm{R}_{\{a\}}([\![\varphi_1]\!]^{\mathfrak{M}^a}) \subseteq \mathrm{R}_\pi([\![\varphi_1]\!]^{\mathfrak{M}^a})$. En cualquier caso, se tiene que $\mathfrak{M} \models [!a]\mathsf{Kh}_i(\varphi_1, \varphi_2)$.
- $\mathfrak{M} \models \mathsf{A}([!a]\varphi_1 \rightarrow (\langle a\rangle\top \wedge [a][!a]\varphi_2))$: Notar que a cumple la condición de ejecutabilidad fuerte y alcanza sólo estados que satisfacen φ. Lo que queda es demostrar que $\{a\} \in \mathrm{U}'(i)$. Sin embargo, esto es sencillo de verificar dado que $\mathrm{U}'(i) = (\mathrm{U}(i) \setminus \{\pi\}) \cup \{\{a\}\}$ si existe $\pi \in \mathrm{U}(i)$ tal que $a \in \pi$, o $\mathrm{U}'(i) = \mathrm{U}(i) \cup \{\{a\}\}$ en caso contrario. Por lo tanto, $\mathfrak{M} \models [!a]\mathsf{Kh}_i(\varphi_1, \varphi_2)$.

□

Ejemplo 9.3. Considerando el LTS^U $\mathfrak{M}$ del Ejemplo 9.1, se tiene que la siguiente instancia de RKh se cumple:

$$\mathfrak{M}, w_1 \models [!a]\mathsf{Kh}_i(p, r) \leftrightarrow (\mathsf{Kh}_i([!a]p, [!a]r) \vee \mathsf{A}([!a]p \rightarrow (\langle a\rangle\top \wedge [a][!a]r))).$$

Si bien $\mathfrak{M}, w_1 \not\models \mathsf{Kh}(p, r)$, se tiene que $\mathfrak{M}, w_1 \models [!a]\mathsf{Kh}(p, r)$. Por lo tanto,

$$\mathfrak{M}, w_1 \models \mathsf{Kh}_i([!a]p, [!a]r) \vee \mathsf{A}([!a]p \to (\langle a \rangle \top \wedge [a][!a]r)).$$

Como $[!a]$ no influye en las variables proposicionales (RAtom), $\mathfrak{M}, w_1 \not\models \mathsf{Kh}_i([!a]p, [!a]r)$ y se tiene que

$$\mathfrak{M}, w_1 \models \mathsf{A}([!a]p \to (\langle a \rangle \top \wedge [a][!a]r)).$$

Quitando $[!a]$ (por RAtom),

$$\mathfrak{M}, w_1 \models \mathsf{A}(p \to (\langle a \rangle \top \wedge [a]r)).$$

Luego, $\mathfrak{M}, w_1 \models \langle a \rangle \top \wedge [a]r$. Esto describe que a es fuertemente ejecutable en el estado w_1 y, desde este, se alcanza sólo estados en donde se cumple r. □

Dada la correctitud de los axiomas de reducción de la Tabla 9.2, cualquier fórmula en $\mathsf{L}_{\mathsf{Kh}_i,\Box,[!a]}$ es equivalente a otra en $\mathsf{L}_{\mathsf{Kh}_i,\Box}$.

Corolario 9.2. $\mathsf{L}_{\mathsf{Kh}_i,\Box} \equiv \mathsf{L}_{\mathsf{Kh}_i,\Box,[!a]}$.

Además, la axiomatización $\mathcal{L}_{\mathsf{Kh}_i,\Box} + \mathcal{L}_{[!a]}$ es correcta y fuertemente completa con respecto a la clase de todos los LTS^Us.

Teorema 9.3. *El sistema axiomático $\mathcal{L}_{\mathsf{Kh}_i,\Box} + \mathcal{L}_{[!a]}$ para $\mathsf{L}_{\mathsf{Kh}_i,\Box,[!a]}$ es correcto y fuertemente completo con respecto a la clase de todos los modelos.*

Demostración. Análoga a la del Teorema 8.2 usando el Lema 9.2 y el Teorema 5.1. □

Finalmente, como la eliminación de modalidades dinámicas toma un número finito de pasos, usando el Corolario 9.1, se tiene lo siguiente.

Corolario 9.3. $\mathsf{L}_{\mathsf{Kh}_i,\Box,[!a]}$ *es decidible.*

9.5. Refinamiento con planes arbitrarios

La lógica introducida en la Sección 9.4 admite una axiomatización correcta y completa sobre la clase de todos los modelos, una propiedad ausente en las lógicas introducidas en el Capítulo 8. Además, dada la limitación de trabajar con acciones atómicas, es lo suficientemente simple para comprender las intuiciones generales. Sin embargo, en esta sección mostraremos que las mismas ideas pueden utilizarse si admitimos planes arbitrarios en estos nuevos anuncios. A continuación se presenta la sintaxis y la semántica de la nueva lógica.

Definición 9.11. Las fórmulas de la lógica $\mathsf{L}_{\mathsf{Kh}_i,\Box,[!\sigma]}$ son las dadas por la siguiente gramática:

$$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid \mathsf{Kh}_i(\varphi, \varphi) \mid [a]\varphi \mid [!\sigma]\varphi,$$

con $p \in \mathsf{Prop}$, $i \in \mathsf{Agt}$, $a \in \mathsf{Act}$ y $\sigma \in \mathsf{Act}^*$. Las fórmulas de la forma $[!\sigma]\varphi$ se leen como *"después de anunciar que σ es distinguible de cualquier otro plan, φ se cumple".*

Definición 9.12. Sean $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \mathrm{U}, \mathrm{V} \rangle$ un LTS^U, $\sigma \in \mathsf{Act}^*$ y φ una fórmula de $\mathsf{L}_{\mathsf{Kh}_i,\Box,[!\sigma]}$,

$$\mathfrak{M}, w \models [!\sigma]\varphi \quad \Leftrightarrow \quad \mathfrak{M}^\sigma, w \models \varphi,$$

donde $\mathfrak{M}^\sigma = \langle \mathrm{S}, \mathrm{R}, \mathrm{U}', \mathrm{V} \rangle$, con:

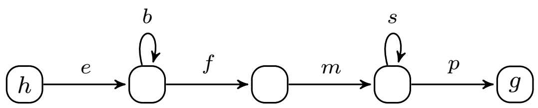
- $\mathrm{U}'(i) = (\mathrm{U}(i) \setminus \{\pi\}) \cup \{\{\sigma\}\}$ si existe $\pi \in \mathrm{U}(i)$ tal que $\sigma \in \pi$,
- $\mathrm{U}'(i) = \mathrm{U}(i) \cup \{\{\sigma\}\}$ en caso contrario.

El modelo refinado $\mathfrak{M}^\sigma$ es esencialmente como en la Definición 9.10 pero para planes arbitrarios. Esta generalización permite hablar del Ejemplo 8.1.

Ejemplo 9.4. Sea $\mathfrak{M}$ el LTS^U del Ejemplo 8.1.

RAtom	$\vdash [!\sigma]p \leftrightarrow p$
R$\neg$	$\vdash [!\sigma]\neg\varphi_1 \leftrightarrow \neg[!\sigma]\varphi_1$
R$\vee$	$\vdash [!\sigma](\varphi_1 \vee \varphi_2) \leftrightarrow ([!\sigma]\varphi_1 \vee [!\sigma]\varphi_2)$
R$\square$	$\vdash [!\sigma][a]\varphi_1 \leftrightarrow [a][!\sigma]\varphi_1$
RKh	$\vdash [!\sigma]\mathsf{Kh}_i(\varphi_1, \varphi_2) \leftrightarrow (\mathsf{Kh}_i([!\sigma]\varphi_1, [!\sigma]\varphi_2) \vee \mathsf{A}([!\sigma]\varphi_1 \rightarrow P(\sigma, [!\sigma]\varphi_2)))$
RE$_{[!]}$	De $\vdash \varphi \leftrightarrow \psi$ se deriva $\vdash [!\sigma]\varphi \leftrightarrow [!\sigma]\psi$

Tabla 9.3: Axiomas de reducción $\mathcal{L}_{[!\sigma]}$.

$\mathfrak{M}$: 

$\mathrm{U}(i) = \{\{ebfmsp\}\}$
$\mathrm{U}(j) = \{\{ebfmsp, ebmfsp\}\}$

Supongamos que un profesor de cocina enseña (anuncia) a ambos agentes que el plan *ebmfsp* es distinguible del resto. Luego cada agente es capaz de distinguir el plan *ebmfsp* de los demás. Más aún, después del anuncio, el agente j sabe cómo hacer una buena torta, descartando el plan incorrecto en el proceso. Por lo tanto, $\mathfrak{M} \models [!ebfmsp]\mathsf{Kh}_j(h, g)$. $\square$

Así como sucedía en la Proposición 9.8, la modalidad es normal y su dual es la misma modalidad.

Proposición 9.9. *Las siguientes propiedades se cumplen:*

1. $\models [!\sigma]\varphi \leftrightarrow \neg[!\sigma]\neg\varphi$.
2. $\models [!\sigma](\varphi_1 \rightarrow \varphi_2) \rightarrow ([!\sigma]\varphi_1 \rightarrow [!\sigma]\varphi_2)$
3. *Si* $\models \varphi$, *entonces* $\models [!\sigma]\varphi$

Demostración. Análoga a la de las Proposiciones 8.2 y 8.4. $\square$

Sistema axiomático

A la hora de deducir una axiomatización completa para $\mathsf{L}_{\mathsf{Kh}_i,\square,[!\sigma]}$, la eliminación de la modalidad dinámica $[!\sigma]$ en presencia de una modalidad Kh_i necesita un poco más de trabajo a comparación de la lógica con acciones atómicas. El caso en el que el anuncio no afecta el conocimiento del agente es directo. Sin embargo, para el otro caso, es necesario describir que para todo estado que cumpla la precondición después del anuncio de σ ($[!\sigma]\varphi_1$), suceda lo siguiente: ***(1)*** σ es fuertemente ejecutable en ese estado, y ***(2)*** después de su ejecución, la postcondición se cumple ($[!\sigma]\varphi_2$).

Estas propiedades son expresadas por el predicado $P(\sigma, [!\sigma]\varphi_2)$, descripto a continuación. Notar que P puede ser definido dado que se ha extendido la expresividad de la lógica base, permitiendo modalidades de acciones básicas.

Lema 9.3. *Sea* $\beta \in \mathsf{Act}^*$, *se define la extensión de* $[a]$ *para planes,* $[\beta]$, *de la forma*

$$[\beta]\psi = \begin{cases} \psi & \beta = \epsilon \\ [\beta[1]] \ldots [\beta[|\beta|]]\psi & \beta \neq \epsilon \end{cases}$$

Con esto, P *se define inductivamente de la siguiente manera:*

$$\begin{array}{rcl} P(\epsilon, \psi) & = & \psi \\ P(\alpha, \psi) & = & (\bigwedge_{k=0}^{|\alpha|-1}([\alpha_k]\langle\alpha[k+1]\rangle\top) \wedge [\alpha]\psi). \end{array}$$

Luego, los axiomas de reducción de la Tabla 9.3 son válidos.

Demostración. Dado que la semántica de $[!\sigma]$ sólo modifica la perspectiva de los agentes (U), los axiomas de reducción RAtom, R$\neg$, R$\vee$ y R$\square$ se cumplen usando la definición de $\models$ (la demostración es similar a la del Lema 8.3). Por otro lado, RE$_{[!]}$ se cumple utilizando la normalidad de

$[!\sigma]$ demostrada en la Proposición 9.9. Por lo tanto, nos queda mostrar que RKh es válida. Sea $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \mathrm{U}, \mathrm{V} \rangle$ un LTS^U.

($\Rightarrow$): $\mathfrak{M} \models [!\sigma]\mathsf{Kh}_i(\varphi_1, \varphi_2)$ sii $\mathfrak{M}^\sigma \models \mathsf{Kh}_i(\varphi_1, \varphi_2)$ sii existe $\pi' \in \mathrm{U}'(i)$ tal que $[\![\varphi_1]\!]^{\mathfrak{M}^\sigma} \subseteq \mathrm{SE}^{\mathfrak{M}^\sigma}(\pi')$ y $\mathrm{R}_{\pi'}([\![\varphi_1]\!]^{\mathfrak{M}^\sigma}) \subseteq [\![\varphi_2]\!]^{\mathfrak{M}^\sigma}$. Usando que $[\![\psi]\!]^{\mathfrak{M}^\sigma} = [\![[!\sigma]\psi]\!]^{\mathfrak{M}}$ y $\mathrm{SE}^{\mathfrak{M}^\sigma}(\pi') = \mathrm{SE}^{\mathfrak{M}}(\pi')$ (las relaciones no cambian), esto es equivalente a que exista $\pi' \in \mathrm{U}'(i)$ tal que $[\![[!\sigma]\varphi_1]\!]^{\mathfrak{M}} \subseteq \mathrm{SE}^{\mathfrak{M}}(\pi')$ y $\mathrm{R}_{\pi'}([\![[!\sigma]\varphi_1]\!]^{\mathfrak{M}}) \subseteq [\![[!\sigma]\varphi_2]\!]^{\mathfrak{M}}$. Se consideran dos casos:

- Si $\pi' \neq \{\sigma\}$, entonces $\pi \in \mathrm{U}(i)$ y se tiene que $\mathfrak{M} \models \mathsf{Kh}_i([!\sigma]\varphi_1, [!\sigma]\varphi_2)$.
- Si $\pi' = \{\sigma\}$, entonces $[\![[!\sigma]\varphi_1]\!]^{\mathfrak{M}} \subseteq \mathrm{SE}^{\mathfrak{M}}(\sigma)$ y $\mathrm{R}_\sigma([\![[!\sigma]\varphi_1]\!]^{\mathfrak{M}}) \subseteq [\![[!\sigma]\varphi_2]\!]^{\mathfrak{M}}$. De la primera parte, usando la definición de ejecutabilidad fuerte, se puede derivar que para todo $w \in [\![[!\sigma]\varphi_1]\!]^{\mathfrak{M}}$, $k \in [0 \ldots |\sigma| - 1]$ se tiene que $\mathfrak{M}, w \models [\sigma_k]\langle\sigma[k+1]\rangle\top$. De la segunda parte, se llega a que para todo $w \in [\![[!\sigma]\varphi_1]\!]^{\mathfrak{M}}$, $\mathfrak{M}, w \models [\sigma][!\sigma]\varphi_2$. Por lo tanto, para todo $w \in \mathrm{S}$, $\mathfrak{M}, w \models [!\sigma]\varphi_1 \to P(\sigma, [!\sigma]\varphi_2)$.

($\Leftarrow$): $\mathfrak{M} \models \mathsf{Kh}_i([!\sigma]\varphi_1, [!\sigma]\varphi_2) \vee \mathsf{A}([!\sigma]\varphi_1 \to P(\sigma, [!\sigma]\varphi_2))$, entonces la disyunción se cumple para alguna de las dos situaciones:

- $\mathfrak{M} \models \mathsf{Kh}_i([!\sigma]\varphi_1, [!\sigma]\varphi_2)$: Sea $\pi \in \mathrm{U}(i)$ tal que $[\![[!\sigma]\varphi_1]\!]^{\mathfrak{M}} \subseteq \mathrm{SE}^{\mathfrak{M}}(\pi)$ y $\mathrm{R}_\pi([\![[!\sigma]\varphi_1]\!]^{\mathfrak{M}}) \subseteq [\![[!\sigma]\varphi_2]\!]^{\mathfrak{M}}$. Como se mencionó anteriormente, esto es equivalente a $[\![\varphi_1]\!]^{\mathfrak{M}^\sigma} \subseteq \mathrm{SE}^{\mathfrak{M}^\sigma}(\pi)$ y $\mathrm{R}_\pi([\![\varphi_1]\!]^{\mathfrak{M}^\sigma}) \subseteq [\![\varphi_2]\!]^{\mathfrak{M}^\sigma}$. Si $\sigma \notin \pi$, entonces $\pi \in \mathrm{U}'(i)$. Si $\sigma \in \pi$, entonces $\{\sigma\} \in \mathrm{U}'(i)$, y $\mathrm{SE}^{\mathfrak{M}^\sigma}(\pi) \subseteq \mathrm{SE}^{\mathfrak{M}^\sigma}(\{\sigma\})$ y $\mathrm{R}_{\{\sigma\}}([\![\varphi_1]\!]^{\mathfrak{M}^\sigma}) \subseteq \mathrm{R}_\pi([\![\varphi_1]\!]^{\mathfrak{M}^\sigma})$. En cualquier caso, se tiene que $\mathfrak{M} \models [!\sigma]\mathsf{Kh}_i(\varphi_1, \varphi_2)$.
- $\mathfrak{M} \models \mathsf{A}([!\sigma]\varphi_1 \to P(\sigma, [!\sigma]\varphi_2))$: Notar que π cumple la condición de ejecutabilidad fuerte y alcanza sólo estados que satisfacen φ. Lo que queda por demostrar es que $\{\sigma\} \in \mathrm{U}'(i)$. Sin embargo, esto es sencillo de verificar dado que $\mathrm{U}'(i) = (\mathrm{U}(i) \setminus \{\pi\}) \cup \{\{\sigma\}\}$ si existe $\pi \in \mathrm{U}(i)$ tal que $\sigma \in \pi$, o $\mathrm{U}'(i) = \mathrm{U}(i) \cup \{\{a\}\}$ en caso contrario. Por lo tanto, $\mathfrak{M} \models [!a]\mathsf{Kh}_i(\varphi_1, \varphi_2)$.

$\square$

Para finalizar esta seccón, análogamente al Teorema 9.3, y a los Corolarios 9.2 y 9.3, se tienen los teoremas de correctitud y completitud fuerte, y los corolarios de equivalencia y decidibilidad para $\mathsf{L}_{\mathsf{Kh}_i,\Box,[!\sigma]}$.

Corolario 9.4. $\mathsf{L}_{\mathsf{Kh}_i,\Box} \equiv \mathsf{L}_{\mathsf{Kh}_i,\Box,[!\sigma]}$.

Teorema 9.4. *El sistema axiomático $\mathcal{L}_{\mathsf{Kh}_i,\Box} + \mathcal{L}_{[!\sigma]}$ para $\mathsf{L}_{\mathsf{Kh}_i,\Box,[!\sigma]}$ es correcto y fuertemente completo con respecto a la clase de todos los modelos.*

Demostración. Análoga a la del Teorema 8.2 usando el Lema 9.3 y el Teorema 5.1. $\square$

Corolario 9.5. $\mathsf{L}_{\mathsf{Kh}_i,\Box,[!\sigma]}$ *es decidible.*

9.6. Refinamiento con planes semi privado

Finalmente, se considera un último tipo de refinamiento. Hasta ahora las modalidades dinámicas introducidas producen actualizaciones "públicas". Esto quiere decir que la información que es revelada es accesible para todos los agentes. Algo similar a lo que se sucede con los operadores de anuncios públicos [91]. Sin embargo, otros tipos de comunicación son posibles. Como por ejemplo, comunicaciones semi privadas o privadas, que se obtienen de, por ejemplo, "action models" o modelos de acción [19], donde la nueva información se hace disponible sólo a un grupo de agentes. En esta sección, se adaptará la modalidad de refinamiento de planes de la Definición 9.9 para producir actualizaciones semi privadas.

Definición 9.13. Las fórmulas de la lógica $\mathsf{L}_{\mathsf{Kh}_i,\Box,[!\sigma,i]}$ son las dadas por la siguiente gramática

$$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid \mathsf{Kh}_i(\varphi, \varphi) \mid [a]\varphi \mid [!\sigma, i]\varphi,$$

con $p \in \mathsf{Prop}$, $i \in \mathsf{Agt}$, $a \in \mathsf{Act}$ y $\sigma \in \mathsf{Act}^*$. Las fórmulas de la forma $[!\sigma, i]\varphi$ se leen como *"después de anunciar al agente i que el plan σ es distinguible de cualquier otro plan, φ se cumple".*

RAtom	$\vdash [!\sigma, i]p \leftrightarrow p$
R$\neg$	$\vdash [!\sigma, i]\neg\varphi_1 \leftrightarrow \neg[!\sigma, i]\varphi_1$
R$\vee$	$\vdash [!\sigma, i](\varphi_1 \vee \varphi_2) \leftrightarrow ([!\sigma, i]\varphi_1 \vee [!\sigma, i]\varphi_2)$
R$\Box$	$\vdash [!\sigma, i][a]\varphi_1 \leftrightarrow [a][!\sigma, i]\varphi_1$
RKh_i	$\vdash [!\sigma, i]\mathsf{Kh}_i(\varphi_1, \varphi_2) \leftrightarrow (\mathsf{Kh}_i([!\sigma, i]\varphi_1, [!\sigma, i]\varphi_2) \vee \mathsf{A}([!\sigma, i]\varphi_1 \rightarrow P(\sigma, [!\sigma, i]\varphi_2)))$
RKh_j	$\vdash [!\sigma, i]\mathsf{Kh}_j(\varphi_1, \varphi_2) \leftrightarrow \mathsf{Kh}_i([!\sigma, i]\varphi_1, [!\sigma, i]\varphi_2)$
$\mathsf{RE}_{[!]}$	De $\vdash \varphi \leftrightarrow \psi$ se deriva $\vdash [!\sigma, i]\varphi \leftrightarrow [!\sigma, i]\psi$

Tabla 9.4: Axiomas de reducción $\mathcal{L}_{[!\sigma,i]}$.

Definición 9.14. Sean $\mathfrak{M} = \langle \mathrm{S}, \mathrm{R}, \mathrm{U}, \mathrm{V} \rangle$ un LTS^U, $\sigma \in \mathsf{Act}^*$ y φ una fórmula de $\mathsf{L}_{\mathsf{Kh}_i,\Box,[!\sigma,i]}$,

$$\mathfrak{M}, w \models [!\sigma, i]\varphi \quad \Leftrightarrow \quad \mathfrak{M}_i^\sigma, w \models \varphi,$$

donde $\mathfrak{M}_i^\sigma = \langle \mathrm{S}, \mathrm{R}, \mathrm{U}', \mathrm{V} \rangle$, con:

- $\mathrm{U}'(j) = \mathrm{U}(j)$ si $j \neq i$,
- $\mathrm{U}'(i) = (\mathrm{U}(i) \setminus \{\pi\}) \cup \{\{\sigma\}\}$ si existe $\pi \in \mathrm{U}(i)$ tal que $\sigma \in \pi$,
- $\mathrm{U}'(i) = \mathrm{U}(i) \cup \{\{\sigma\}\}$ en caso contrario.

Ejemplo 9.5. Supongamos que en el Ejemplo 8.1, tenemos el cocinero experto i, pero tenemos dos inexpertos j y k, con $\mathrm{U}(j) = \mathrm{U}(k) = \{\{\mathit{ebfmsp}, \mathit{ebmfsp}\}\}$. Supongamos que el agente j se suscribe a un curso dictado por el agente i, mientras que el agente k no. Entonces la información de que *ebfmsp* es el curso de acción correcto es revelada únicamente a j, dando a que $\mathfrak{M} \models [!ebfmsp, j]\mathsf{Kh}_j(h, g)$ mientras que $\mathfrak{M} \models [!ebfmsp, j]\neg\mathsf{Kh}_k(h, g)$. $\Box$

Nuevamente, esta modalidad es normal y su dual es la misma.

Proposición 9.10. *Las siguientes propiedades se cumplen:*

1. $\models [!\sigma, i]\varphi \leftrightarrow \neg[!\sigma, i]\neg\varphi$.
2. $\models [!\sigma, i](\varphi_1 \rightarrow \varphi_2) \rightarrow ([!\sigma, i]\varphi_1 \rightarrow [!\sigma, i]\varphi_2)$
3. *Si* $\models \varphi$, *entonces* $\models [!\sigma, i]\varphi$

Los axiomas de reducción de esta lógica con anuncios semi privados descriptos en la Tabla 9.4, y similares a la versión "pública" en la Tabla 9.2, son válidos.

Lema 9.4. *Sea P la fórmula definida en el Lema 9.3. Entonces, los axiomas de la Tabla 9.4 son válidos.*

Demostración. Análoga a la del Lema 9.3 usando la Proposición 9.10. $\Box$

Una vez más, se obtienen los teoremas de correctitud y completitud fuerte, y los corolarios de equivalencia y decidibilidad para $\mathsf{L}_{\mathsf{Kh}_i,\Box,[!\sigma,i]}$.

Corolario 9.6. $\mathsf{L}_{\mathsf{Kh}_i,\Box} \equiv \mathsf{L}_{\mathsf{Kh}_i,\Box,[!\sigma,i]}$.

Teorema 9.5. *El sistema axiomático $\mathcal{L}_{\mathsf{Kh}_i,\Box} + \mathcal{L}_{[!\sigma,i]}$ para $\mathsf{L}_{\mathsf{Kh}_i,\Box,[!\sigma,i]}$ es correcto y fuertemente completo con respecto a la clase de todos los modelos.*

Demostración. Análoga a la del Teorema 8.2 usando el Lema 9.2 y el Teorema 5.1. $\Box$

Corolario 9.7. $\mathsf{L}_{\mathsf{Kh}_i,\Box,[!\sigma,i]}$ *es decidible.*

El presente capítulo, en conjunto con el Capítulo 8, nos da cuenta de la versatilidad de la semántica basada en LTS^U a la hora de expresar y definir las propiedades deseadas. En particular, nos enfocamos en diferentes formas de comunicación de información mediante operadores dinámicos que pueden producir cambios en el conocimiento de los agentes. Podemos concluir del Capítulo 9 que, pese a las dificultades que aparecen a la hora de axiomatizar modalidades dinámicas, es posible lidiar con ellas agregando expresividad al lenguaje básico, siendo en este caso operadores modales básicos $[a]$. Naturalmente existen muchas otras alternativas, y surge la pregunta de qué expresividad necesitamos para axiomatizar operadores más generales. Nuestro trabajo en este capítulo es el punto de inicio para conseguir este tipo de resultados en el futuro.

Capítulo 10

Una interpretación deóntica: cumplir conscientemente con las normas

> "Todo me está permitido", dicen algunos. Sí, pero no todo es conveniente.
>
> 1 Corintios 6, 12a

Los sistemas normativos [46] se encuentran en múltiples disciplinas como el razonamiento legal, las ciencias de la computación, la representación del conocimiento, la filosofía, entre otros. Por lo tanto, analizar de manera rigurosa y lógica sobre estos sistemas resulta de gran interés. Uno de los enfoques lógicos más destacados para el razonamiento normativo son las lógicas deónticas [20, 68, 113]. Estas lógicas son formalismos diseñados para describir y razonar sobre escenarios que involucran normas y conceptos relacionados [46, 47], y pueden ser usados para determinar, por ejemplo, si dichos escenarios son libres de contradicciones o paradojas.

Normalmente, las lógicas deónticas poseen operadores para hablar sobre obligaciones, permisos y prohibiciones de los agentes involucrados. Se ha discutido, por ejemplo en [113, 114], que estos operadores deben expresar propiedades sobre las acciones ejecutadas por los agentes, en vez describir información proposicional o factual. Las lógicas deónticas basadas en la primera característica son comúnmente conocidas como lógicas deónticas de *deber hacer* [4], en contraposición con las que se basan en la segunda característica, denominadas lógicas deónticas de *deber ser*, que están más enfocadas en describir prescripciones sobre proposiciones. En gran medida, el enfoque de las lógicas deónticas de deber hacer, es el estado normativo de las acciones en un momento específico. Esto ignora, por ejemplo, las circunstancias que conllevan a que un agente alcance un determinado estado, y el *curso de acción* que un agente toma para lograr un objetivo. Para tratar estas situaciones, el campo de la lógica deóntica incorpora ideas del área de razonamiento estratégico (ver, por ejemplo, [31, 64]). En una lógica deóntica de deber hacer enriquecida con características relativas a las estrategias que un agente puede ejecutar, la descripción de una situación normativa considera no sólo el estado de un conjunto de normas, sino que también el comportamiento de los agentes con respecto a su entorno.

En una lógica deóntica de deber hacer, el razonamiento estratégico usualmente se consigue incorporando maneras de describir secuencias de acciones o planes sobre una dimensión temporal. Dicha dimensión es entonces usada para modelar y estudiar un sistema normativo particular. Varios marcos teóricos de este tipo extienden la lógica STIT [22] con operadores temporales, por ejemplo en [27, 28]. Por ello, los operadores temporales de "necesidad histórica", junto con un operador de "saber qué", son incorporados a la lógica STIT. El sistema lógico resultante termina adaptándose a la noción de *hacer conscientemente*, cuyo objetivo es caracterizar diferentes modos de responsabilidad para un agente que infringe una norma (un concepto conocido como "mente culpable" o "mens rea"). El trabajo en [29] toma un enfoque similar pero incorpora un operador de "intención" para representar distintos niveles de culpabilidad. Finalmente, en [78] se investiga la lógica T-STIT, una extensión de la lógica STIT con operadores de pasado y futuro, y con un operador de grupos para coaliciones involucrando a todos los agentes. La lógica obtenida se utiliza para modelar conceptos normativos como la obligación de logro y el compromiso (para más detalles ver [64, Capítulo 7]).

Los marcos teóricos mencionados son altamente expresivos dado que combinan varios operadores de diferente naturaleza. Sin embargo, la formalización de un concepto deóntico particular termina usando sólo un fragmento de estas lógicas, con una expresividad alta que afecta negativamente el comportamiento computacional de la lógica. Existen ejemplos de fragmentos de la lógica STIT cuyo problema de satisfabilidad es indecidible [15, 60], otros que se encuentran en ExpTime, NExpTime, y, en el mejor de los casos, NP pero removiendo la parte temporal o limitando las negaciones anidadas.

Este capítulo tiene un enfoque distinto a estas propuestas basadas en lógicas STIT y se alinea con lo que se ha visto a lo largo de esta tesis. En vez de tomar operadores existentes en la literatura y tratar de dar una formalización de ciertas situaciones, se tomará un concepto deóntico relevante que involucre un razonamiento estratégico específico y se definirá una lógica que formaliza tal concepto. En este sentido, trabajaremos con la noción de *cumplir conscientemente* con las normas. Para esto, se definirá una lógica que represente este concepto de manera natural y que tenga un buen comportamiento computacional, tomando como inspiración la lógica basada en incertidumbre introducida en el Capítulo 4 y la terminología de [23], en la cual consideraremos tres componentes importantes para definir esta noción de *cumplir conscientemente*: ***(1)*** las *habilidades* de los agentes, es decir, lo que los agentes son *capaces de hacer* (todas las estrategias o cursos de acción posibles que pueden tomar para cumplir un objetivo); ***(2)*** el conjunto de *normas*, que establecen lo que los agentes *deben* hacer (se expresan como un conjunto legal de cursos de acción); y ***(3)*** las *responsabilidades* de los agentes, dadas por el *juicio* de cada agente (permite determinar la consciencia de cada agente a la hora de cumplir o no con una norma). Este enfoque fue introducido y desarrollado en [5] y resulta natural en las lógicas modales. De la misma manera que la lógica modal básica posee diferentes interpretaciones (deóntica, epistémica, temporal, etcétera), dependiendo de los axiomas y las clases de modelos consideradas, para el caso de "saber cómo" podemos aplicar un razonamiento similar, y estudiar la semántica introducida en otros contextos, en este caso, en uno deóntico.

10.1. Motivación

Consideremos una situación en la que se encuentran las tres componentes mencionadas anteriormente: un plan de evacuación de un edificio. Usualmente, estos planes son documentos escritos detallando las acciones que deben llevarse a cabo en caso de un incendio y las disposiciones para llamar a los bomberos. La Figura 10.1 muestra un procedimiento de emergencia arquetípico y contiene algunas normas importantes que se deben cumplir. Por ejemplo, en caso de fuego, humo o una explosión, cada persona debe seguir el siguiente curso de acción: activar la alarma, ir a un lugar seguro, y desde allí llamar a los bomberos. Más aún, cada persona debe evacuar teniendo en cuenta riesgos básicos y otras precauciones como cerrar las puertas detrás o, en caso de no poder evacuar, quedarse y bloquear las aberturas. Este procedimiento asume que las personas tienen el conocimiento de la distribución de las instalaciones, la capacidad de identificar rutas de escape y la posibilidad de desalojar las instalaciones utilizando sólo escaleras y/o rampas (los ascensores están excluidos ya que pueden fallar por cortes de electricidad). Además, el procedimiento establece que cada persona debe mantener la calma en todo momento, aún si toma un curso de acción equivocado (por ejemplo, tomar el ascensor).

Con esto, se pueden distinguir las tres componentes mencionados anteriormente: ***(1)*** las habilidades son las posibles acciones o planes que pueden ejecutar los agentes, ya sea activar la alarma, usar las escaleras, la rampa, o incluso el ascensor aunque no estuviera permitido; ***(2)*** las normas son los planes "legales" y reflejan los cursos de acción que están permitidos como el plan "activar la alarma, tomar las escaleras (o la rampa) y llamar al 100"; y ***(3)*** las responsabilidades que capturan el propio juicio de cada agente para cumplir con las normas.

Este último punto requiere más detalle. Desde la perspectiva de un agente particular, pueden existir ciertos cursos de acción que son indistinguibles de otros. Por ejemplo, un agente i puede considerar que salir del edificio usando las escaleras o la rampa está bien, mientras que evita usar el ascensor, siguiendo con el procedimiento normado. Mientras tanto, un agente j que ignora el plan de evacuación, puede considerar que salir del edificio sin importar la forma es lo importante. En nuestro sistema formal deseamos expresar que el agente i cumple conscientemente con las

PROCEDIMIENTO DE EMERGENCIA

– FUEGO – HUMO – EXPLOSIÓN	MANTENER LA CALMA ACTIVAR LA ALARMA, DESDE UN LUGAR SEGURO LLAMAR AL 100 (BOMBEROS)

AL SONAR LA ALARMA, APAGAR GAS Y ELECTRICIDAD

Al evacuar: cerrar las puertas atrás, usar sólo las escaleras y las rampas, nunca el ascensor.
Si no es seguro evacuar: cerrar la puerta, bloquear aberturas, mantenerse cerca de las ventanas.

Figura 10.1: Plan de evacuación en caso de incendio

normas, mientras que j no. Más aún, algunos agentes pueden no ser conscientes de determinados cursos de acción posibles. Por ejemplo, la ubicación al desconocer una salida de emergencia.

En las siguientes secciones, estos conceptos de habilidades, normas y cumplimiento a consciencia serán presentados formalmente. Para ello, será necesario generalizar la semántica que se estuvo trabajando desde los Capítulos 3 y 4, tomando la estructura del LTS como el componente que representa las habilidades disponibles para todos los agentes, un conjunto de planes que establece los cursos de acción normativos, y finalmente, una relación de indistinguibilidad entre planes que captura la percepción de cada agente sobre las acciones que puede tomar, tal como se modelaba la percepción del agente en nuestro entorno de saber cómo. Luego, se introducirán las modalidades correspondientes que nos permitan hablar de las propiedades deónticas sobre este tipo de modelos: la modalidad $\mathsf{N}(\psi, \varphi)$ que establece que *"en cualquier situación en la cual ψ se cumple, es posible conseguir el objetivo φ de acuerdo con las normas"*, y la modalidad $\mathsf{Kc}_i(\psi, \varphi)$, para cada agente i, que establece que *"en cualquier situación en la que ψ se cumple, el agente i cumple conscientemente con φ"*. Investigaremos las propiedades lógicas de esta lógica deóntica, proveyendo un sistema axiomático correcto y fuertemente completo, y demostrando que sus problemas de model checking se encuentra en P y que satisfabilidad es NP-completo. Luego, extenderemos el lenguaje con una modalidad $\mathsf{S}(\psi, \varphi)$ que nos permite expresar propiedades sobre las habilidades de los agentes independientemente de las normas establecidas y su percepción, y su interacción con las modalidades anteriores. La misma se lee como *"en cualquier situación en la cual ψ se cumple, es posible conseguir el objetivo φ"*. Al capturar las habilidades en el lenguaje, esto permite distinguir lo que los agentes *pueden hacer* de lo que *deben hacer de acuerdo a las normas*. Para esta extensión se tiene una axiomatización correcta y fuertemente completa.

10.2. Lógica deóntica

En esta sección se introduce la sintaxis y la semántica de la lógica deóntica con los operadores normativos N y Kc_i, denotada como $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N}}$.

Definición 10.1. Sean Prop y Act conjuntos numerables de símbolos proposicionales y de acciones básicas respectivamente, y sea Agt un conjunto no vacío de nombres de agentes. Las fórmulas de $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N}}$ son dadas por la siguiente gramática

$$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid \mathsf{N}(\varphi, \varphi) \mid \mathsf{Kc}_i(\varphi, \varphi),$$

con $p \in \mathsf{Prop}$ e $i \in \mathsf{Agt}$. Las constantes y demás conectores booleanos $\bot$, $\top$, $\wedge$, $\rightarrow$, y $\leftrightarrow$ se definen de la forma usual. Intuitivamente, la fórmula $\mathsf{N}(\psi, \varphi)$ se lee como *"existe un curso de acción normativo que cumple φ dado ψ"*, y la fórmula $\mathsf{Kc}_i(\psi, \varphi)$ se lee como *"el agente i cumple conscientemente con φ dado ψ"*.

Más adelante, como sucede en las Proposiciones 3.1 y 4.1, mostraremos que las modalidades universal A y existencial E [51] también son definibles en esta lógica (Proposición 10.1).

En el Capítulo 4, la introducción de una relación de indistinguibilidad permitió separar las habilidades (representadas en el LTS) de la percepción que los agentes tienen sobre estas (respresentada en U). En el caso de $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N}}$, las estructuras que se utilizan en la parte semántica son similares a las de $\mathsf{L}_{\mathsf{Kh}_i}$ (ver las Definiciones 3.4 y 4.2, y la Observación 4.1) y se utilizan los mismos conceptos de acciones básicas, planes y ejecutabilidad fuerte. Sin embargo, existe una diferencia importante. Dado que se quiere modelar si el agente cumple conscientemente o no con las normas (es decir, las responsabilidades), el foco de la relación de indistinguibilidad no es sólo con respecto a las *habilidades*, sino también a las *normas*. Para ello, será necesario introducir una variación de los modelos LTS^U considerando esta dimensión "normativa".

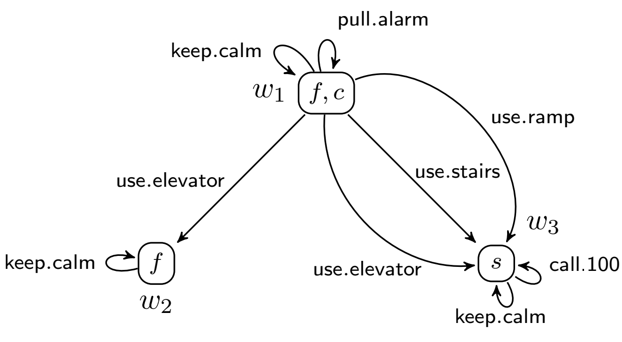
Definición 10.2 (LTS^U Normativo)**.** Un LTS^U *normativo* (NLTS^U) es una tupla $\mathfrak{N} = \langle \mathrm{S}, \mathrm{R}, \mathrm{V}, \mathrm{U}, \mathrm{N}\rangle$ donde:

- $\mathfrak{L} = \langle \mathrm{S}, \mathrm{R}, \mathrm{V}\rangle$ es un LTS;
- $\mathrm{U} : \mathsf{Agt} \to 2^{2^{\mathsf{Act}^*}}$ es una función tal que ***(1)*** $\mathrm{U}(i) \neq \emptyset$, ***(2)*** $\emptyset \in \mathrm{U}(i)$, y ***(3)*** para todo $\pi, \pi' \in \mathrm{U}(i)$, $\pi \neq \pi'$ implica $\pi \cap \pi' = \emptyset$; y
- $\mathrm{N} \subseteq \mathsf{Act}^*$ es un conjunto de planes tal que existe un plan $\sigma \in \mathrm{N}$ fuertemente ejecutable en todos los estados ($\mathrm{SE}(\sigma) = \mathrm{S}$).

El conjunto N en un NLTS^U es denominado como el conjunto de planes normativos. El requerimiento de que tenga al menos un plan fuertemente ejecutable puede entenderse de que "existen normas que siempre pueden cumplirse". Esto toma inspiración de la propiedad deóntica de que "deber" implica "poder" (denominada serialidad en lógica modal). Por otro lado, la definición del conjunto U es casi idéntica a las vistas en la Definición 4.2 y la Observación 4.1. Como siempre, esta representa los cursos de acción que son indistinguibles desde la percepción de cada agente. Es decir, cada $\pi \in \mathrm{U}(i)$ captura los cursos de acción que, desde la perspectiva del agente i, son tan buenos como cualquier otro dentro de este conjunto. Sin embargo, hay una sutil diferencia y es que admitimos el conjunto vacío en cada $\mathrm{U}(i)$ ($\emptyset \in \mathrm{U}(i)$). Esta condición indica que el plan de 'abortar' (es decir un plan que falla siempre) es posible, y que los agentes pueden distinguirlo del resto. Con esto, se define que $\mathrm{SE}(\emptyset) = \emptyset$. Discutiremos más adelante los motivos técnicos para este cambio.

Ejemplo 10.1. Volviendo al ejemplo del procedimiento de emergencia, el mismo se puede representar de manera simplificada con el siguiente LTS. El mismo consta de tres estados: ***(1)*** w_1 representa la situación en la que hay un incendio (representado por el símbolo proposicional f) y el agente posee la capacidad de seguir el protocolo de evacuación (c); ***(2)*** w_2 representa la situación en la que hay un incendio (f), pero no hay capacidad de seguir dicho protocolo ($\neg c$) (por ejemplo, por estar atrapado en el ascensor); y ***(3)*** w_3 representa la situación en la que se llegó a un lugar seguro (s) y se está fuera de peligro del incendio ($\neg f$).

Las acciones básicas consideradas en este caso son las que aparecen en el procedimiento. Estas son $\mathsf{keep.calm}$ (mantener la calma), $\mathsf{pull.alarm}$ (activar la alarma), $\mathsf{call.100}$ (llamar al 100), $\mathsf{use.stairs}$ (usar las escaleras), $\mathsf{use.ramp}$ (usar la rampa) y $\mathsf{use.elevator}$ (usar el ascensor).



En este contexto, los planes recomendados por el procedimiento son

$$\sigma_1 = \mathsf{pull.alarm}; \mathsf{use.stairs}; \mathsf{call.100}$$

$$\sigma_2 = \mathsf{pull.alarm}; \mathsf{use.ramp}; \mathsf{call.100}$$

(los punto y coma son para separar las acciones entre sí) que a su vez son fuertemente ejecutables en el estado w_1 y llegan al estado w_3. Por otro lado, como el procedimiento de emergencia desestima usar el ascensor, en el LTS el plan $\sigma_3 = \mathsf{use.elevator}$ es fuertemente ejecutable en w_1 y puede llevar al estado w_2 o w_3. Si este plan a su vez se lo compone con las acciones de activar la alarma y llamar a los bomberos, se tiene

$$\sigma_4 = \mathsf{pull.alarm}; \mathsf{use.elevator}; \mathsf{call.100}$$

el cual no es fuertemente ejecutable en w_1 dado que, al ejecutar $\mathsf{use.elevator}$, se puede llegar al estado w_2 donde no es posible tomar la acción $\mathsf{call.100}$. Por otro lado, el plan $\sigma_5 = \mathsf{keep.calm}$ es fuertemente ejecutable en todos los estados.

Para modelar los cursos de acción permitidos por el protocolo y que cumplan con la Definición 10.2, es razonable definir $\mathrm{N} = \{\sigma_1, \sigma_2, \sigma_5\}$, siendo el plan σ_5 la norma que cumple el requisito de ser fuertemente ejecutable en todos los estados, y los planes σ_1 y σ_2 las normas de evacuación y notificación a las autoridades.

Supongamos además que tenemos dos agentes i y j. El agente i ha tomado un curso de capacitación en normas de seguridad, mientras que j no. Por lo tanto, i debería saber la diferencia entre usar las escaleras o la rampa y usar el ascensor para evacuar el edificio en caso de un incendio. Por otro lado, para j, todas las formas posibles de salir del edificio no tienen diferencia para él. Esto se representa definiendo $\mathrm{U}(i) = \{\emptyset, \{\sigma_1, \sigma_2\}, \{\sigma_4\}\}$ y $\mathrm{U}(j) = \{\emptyset, \{\sigma_1, \sigma_2, \sigma_4\}\}$ respectivamente. Con el LTS anterior, el conjunto de normas N y la percepción de los agentes de U, se obtiene un NLTS^U. □

Teniendo en cuenta estos elementos, definimos la semántica de $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N}}$. Dicha semántica tiene varias similitudes con las dadas en las Definiciones 3.7 y 4.5 dado que persisten las condiciones de ejecutabilidad fuerte y alcanzabilidad para los operadores N y Kc_i.

Definición 10.3 ($\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N}}$ en NLTS^Us)**.** Sea $\mathfrak{N} = \langle \mathrm{S}, \mathrm{R}, \mathrm{V}, \mathrm{U}, \mathrm{N} \rangle$ un NLTS^U sobre Act, Prop y Agt, y $w \in \mathrm{S}$. La relación de satisfacibilidad $\models$ entre $\mathfrak{N}$, w y las fórmulas de $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N}}$ se define inductivamente de la siguiente manera:

$\mathfrak{N}, w \models p$	sii	$p \in \mathrm{V}(w)$,
$\mathfrak{N}, w \models \neg\varphi$	sii	$\mathfrak{N}, w \not\models \varphi$,
$\mathfrak{N}, w \models \varphi \vee \psi$	sii	$\mathfrak{N}, w \models \varphi$ o $\mathfrak{N}, w \models \psi$,
$\mathfrak{N}, w \models \mathsf{N}(\psi, \varphi)$	sii	existe $\sigma \in \mathrm{N}$ tal que
		(N-1) $[\![\psi]\!]^{\mathfrak{N}} \subseteq \mathrm{SE}(\sigma)$, y **(N-2)** $\mathrm{R}_\sigma([\![\psi]\!]^{\mathfrak{N}}) \subseteq [\![\varphi]\!]^{\mathfrak{N}}$
$\mathfrak{N}, w \models \mathsf{Kc}_i(\psi, \varphi)$	sii	existe $\pi \in \mathrm{U}(i)$ tal que
		(Kc-1) $\pi \subseteq \mathrm{N}$, **(Kc-2)** $[\![\psi]\!]^{\mathfrak{N}} \subseteq \mathrm{SE}(\pi)$, y **(Kc-3)** $\mathrm{R}_\pi([\![\psi]\!]^{\mathfrak{N}}) \subseteq [\![\varphi]\!]^{\mathfrak{N}}$

con $[\![\chi]\!]^{\mathfrak{N}} = \{w \in \mathrm{S} \mid \mathfrak{N}, w \models \chi\}$. Una fórmula φ es satisfacible sii $\mathfrak{N}, w \models \varphi$ para algún $\mathfrak{N}$ y $w \in \mathrm{D}_{\mathfrak{N}}$, y es válida (denotado como $\models \varphi$) sii $\mathfrak{N}, w \models \varphi$ para todo $\mathfrak{N}$ y $w \in \mathrm{D}_{\mathfrak{N}}$. Una fórmula φ se cumple (globalmente) en un NLTS^U $\mathfrak{N}$ ($\mathfrak{N} \models \varphi$) sii $\mathfrak{N}, w \models \varphi$ para todo $w \in \mathrm{D}_{\mathfrak{N}}$. Para $\Psi \subseteq \mathsf{L}_{\mathsf{Kh}}$, $\mathfrak{N}, w \models \Psi$ sii $\mathfrak{N}, w \models \psi$ para todo $\psi \in \Psi$; y $\Psi \models \varphi$ sii $\mathfrak{N}, w \models \Psi$ implica $\mathfrak{N}, w \models \varphi$ para todo $\mathfrak{N}$ y $w \in \mathrm{D}_{\mathfrak{N}}$. El plan σ (respectivamente el conjunto de planes π) que hace verdadero $\mathsf{N}(\psi, \varphi)$ (respectivamente $\mathsf{Kh}_i(\psi, \varphi)$) en $\mathfrak{N}$ es llamado *testigo normativo* (respectivamente, *testigo normativo de i*) para $\mathsf{N}(\psi, \varphi)$ (respectivamente $\mathsf{Kh}_i(\psi, \varphi)$) en $\mathfrak{N}$.

De manera intuitiva, $\mathsf{N}(\psi, \varphi)$ establece que existe un plan normativo que alcanza φ desde ψ. Mientras $\mathsf{Kc}_i(\psi, \varphi)$ establece que existe un conjunto de planes normativos, todos indistinguibles desde la perspectiva del agente i, los cuales cada uno alcanza φ desde ψ. Tomando una lectura normativa del operador Kc_i, se tiene que el agente sabe cómo cumplir φ dada una condición inicial ψ dentro de los límites de la ley, distinguiendo los cursos de acción testigo como normativos de otros que no lo son. Por ello nos referimos a Kc_i como *cumplir conscientemente*.

Ejemplo 10.2. Dado el Ejemplo 10.1, se tiene que $\mathfrak{N}, w_1 \models \mathsf{N}(f \wedge c, s)$, es decir, existe un curso de acción normativo que permite a los agentes llegar a un lugar seguro, desde cualquier estado donde hay un incendio y existe la capacidad de seguir el protocolo, eligiendo como testigo normativo σ_1 ó σ_2. El agente i puede cumplir a consciencia con el protocolo, $\mathfrak{N}, w_1 \models \mathsf{Kc}_i(f \wedge c, s)$, siendo $\{\sigma_1, \sigma_2\}$ el testigo normativo de i. Sin embargo, para el caso del agente j, se tiene que $\mathfrak{N}, w_1 \not\models \mathsf{Kc}_j(f \wedge c, s)$. Esto se debe a que j ***(1)*** confunde los planes normativos σ_1 y σ_2 con σ_4 que no lo es ($\{\sigma_1, \sigma_2, \sigma_4\} \not\subseteq \mathrm{N}$); y ***(2)*** el plan $\emptyset$ falla en todos los estados incluyendo w_1 ($[\![f \wedge c]\!]^{\mathfrak{N}} = \{w_1\} \not\subseteq \mathrm{SE}(\emptyset) = \emptyset$). Finalmente, al tener un plan fuertemente ejecutable en todos los estados, se tiene que $\mathfrak{N}, w_1 \models \mathsf{N}(\psi, \top)$, independientemente de ψ. □

Cómo se mencionó anteriormente, las modalidades universal y existencial son definibles en esta lógica de la forma $\mathsf{A}\varphi := \mathsf{N}(\neg\varphi, \bot)$ y $\mathsf{E}\varphi := \neg\mathsf{A}\neg\varphi$. Sin embargo, dada la globalidad de estas modalidades, existen otras definiciones alternativas como $\mathsf{N}(\neg\varphi, \bot) \vee \bigvee_{i\in\mathsf{Agt}} \mathsf{Kc}_i(\neg\varphi, \bot)$ ó $\mathsf{N}(\neg\varphi, \bot) \wedge \bigwedge_{i\in\mathsf{Agt}} \mathsf{Kc}_i(\neg\varphi, \bot)$. Enfoques similares se han visto en el Capítulo 4 con las Proposiciones 4.1, 4.2 y 4.3. Por el momento, nos quedaremos con la primera definición en términos N únicamente. La siguiente proposición muestra que N captura la semántica de la modalidad universal. La demostración se basa en que el conjunto N nunca es vacío ya que siempre tiene un plan σ tal que $\mathrm{SE}(\sigma) = \mathrm{S}$, un argumento similar al dado en la Proposición 4.1 para U con el operador Kh_i.

Proposición 10.1. *Sean* $\mathfrak{N} = \langle \mathrm{S}, \mathrm{R}, \mathrm{V}, \mathrm{U}, \mathrm{N}\rangle$ *un* NLTS^U, $w \in \mathrm{S}$, *y* φ *una fórmula. Entonces,*

$$\mathfrak{N}, w \models \mathsf{N}(\neg\varphi, \bot) \qquad \textit{sii} \qquad [\![\varphi]\!]^{\mathfrak{N}} = \mathrm{S}.$$

Demostración. La demostración es similar a la vista en la Proposición 4.1 para Kh_i. □

Ejemplo 10.3. Con el modelo del Ejemplo 10.1, $\mathfrak{N}, w \models \mathsf{A}(s \to \neg f)$ (todo lugar que es seguro no está incendiándose) y $\mathfrak{N}, w \models \mathsf{E}f$ (existe una situación en la que hay un incendio). □

Con la sintaxis y semántica definidas, podemos introducir el sistema axiomático para $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N}}$. El mismo contará con similitudes con respecto a los vistos para L_{Kh} (Tabla 3.1), $\mathsf{L}_{\mathsf{Kh}_i}$ (Tabla 5.1) y $\mathsf{L}_{\mathsf{Kh}_i,\Box}$ (Tabla 9.1).

10.3. Axiomatización

En esta sección presentamos una axiomatización correcta y fuertemente completa para $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N}}$ (Tabla 10.1), la cual fue presentada ya en [5]. La misma consta de dos reglas y dos bloques de axiomas, $\mathcal{L}$ y $\mathcal{L}_{\mathrm{NLTS}^U}$, similares a los introducidos en Tabla 5.1, generalizando los axiomas 4KhA, 5KhA y KhA para los operadores Kc_i y N. Por último, se tienen los siguientes axiomas:

1. KcN: Vincula las modalidades Kc_i y N. Todo lo que el agente i cumple conscientemente está regulado por las normas. Este axioma es similar a $\mathsf{A}\Box$ de la Tabla 9.1.

2. DN: Describe la serialidad de N. Desde cualquier situación, existe una forma legal de conseguir un objetivo universalmente válido (atestiguada por el plan fuertemente ejecutable en todos los estados).

3. NKc$\bot$: Vincula las globalidades de A, definida en términos de N, con Kc_i. Esto deviene de la globalidad de ambos operadores. Notar que aquí se hace indispensable la pertenencia del conjunto vacío en $\mathrm{U}(i)$. Dado que es natural contar con esta conexión entre ambas modalidades, para cumplir con tal implicación es necesario admitir siempre un conjunto de planes que trivialmente pertenezcan a N. Podríamos generalizar esta condición para un conjunto arbitrario, pero consideramos que el conjunto vacío representando la acción de abortar tiene sentido en este contexto.

La correctitud de $\mathcal{L}_{\mathsf{Kc}_i,\mathsf{N}}$ es directa dado que cada una de las reglas y axiomas son válidos para modelos arbitrarios. Similarmente a los Capítulos 5 y 9, la demostración de completitud (Teorema 10.1) requiere definir un modelo canónico y demostrar propiedades del mismo. Como en las Definiciones 5.1 y 9.3, se introducen los conceptos de deducción, consistencia y maximalidad.

Definición 10.4. Sea $\Gamma \cup \{\varphi\}$ un conjunto fórmulas de $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N}}$. Las nociones de deducción, $\Gamma \vdash \varphi$, y teorema, $\vdash \varphi$, se definen de la manera usual en el contexto de $\mathcal{L}_{\mathsf{Kc}_i,\mathsf{N}}$. Un conjunto Γ es consistente con $\mathcal{L}_{\mathsf{Kc}_i,\mathsf{N}}$ si $\Gamma \nvdash \bot$, y Γ es maximal consistente con $\mathcal{L}_{\mathsf{Kc}_i,\mathsf{N}}$ si Γ es consistente con $\mathcal{L}_{\mathsf{Kc}_i,\mathsf{N}}$ y para todo $\varphi \notin \Gamma$, $\Gamma \cup \{\varphi\} \vdash \bot$.

En primer lugar, se tienen los siguientes teoremas derivables de la Tabla 10.1.

Proposición 10.2. *Las siguientes fórmulas son derivables usando los axiomas y reglas de* $\mathcal{L}_{\mathsf{Kc}_i,\mathsf{N}}$*:* *(1) Kc$\bot$:* $\vdash \mathsf{Kc}_i(\bot, \bot)$*; (2) KcE:* $\vdash (\mathsf{E}\psi \wedge \mathsf{Kc}_i(\psi, \varphi)) \to \mathsf{E}\varphi$*; (3) NE:* $\vdash (\mathsf{E}\psi \wedge \mathsf{N}(\psi, \varphi)) \to \mathsf{E}\varphi$.

Demostración.

(Kc$\bot$) Por TAUT y NECA, $\vdash \mathsf{A}\top$. Sintácticamente, $\mathsf{A}\top = \mathsf{N}(\neg\top, \bot) = \mathsf{N}(\bot, \bot)$. Usando NKc$\bot$ y MP, se tiene que $\vdash \mathsf{Kc}_i(\bot, \bot)$.

Bloque $\mathcal{L}$:	TAUT	$\vdash \varphi$ para φ una tautología proposicional
	DISTA	$\vdash \mathsf{A}(\varphi \to \psi) \to (\mathsf{A}\varphi \to \mathsf{A}\psi)$
	TA	$\vdash \mathsf{A}\varphi \to \varphi$
	4KcA	$\vdash \mathsf{Kc}_i(\psi, \varphi) \to \mathsf{AKc}_i(\psi, \varphi)$
	5KcA	$\vdash \neg\mathsf{Kc}_i(\psi, \varphi) \to \mathsf{A}\neg\mathsf{Kc}_i(\psi, \varphi)$
	4NA	$\vdash \mathsf{N}(\psi, \varphi) \to \mathsf{AN}(\psi, \varphi)$
	5NA	$\vdash \neg\mathsf{N}(\psi, \varphi) \to \mathsf{A}\neg\mathsf{N}(\psi, \varphi)$
Bloque $\mathcal{L}_{\mathrm{NLTS}^U}$:	KcA	$\vdash (\mathsf{A}(\psi \to \chi) \wedge \mathsf{Kc}_i(\chi, \rho) \wedge \mathsf{A}(\rho \to \varphi)) \to \mathsf{Kc}_i(\psi, \varphi)$
	NA	$\vdash (\mathsf{A}(\psi \to \chi) \wedge \mathsf{N}(\chi, \rho) \wedge \mathsf{A}(\rho \to \varphi)) \to \mathsf{N}(\psi, \varphi)$
	KcN	$\vdash \mathsf{Kc}_i(\psi, \varphi) \to \mathsf{N}(\psi, \varphi)$
	DN	$\vdash \mathsf{N}(\psi, \top)$
	NKc⊥	$\vdash \mathsf{N}(\psi, \bot) \to \mathsf{Kc}_i(\psi, \bot)$
Reglas:	$\dfrac{\varphi \quad (\varphi \to \psi)}{\psi}$ MP	$\dfrac{\varphi}{\mathsf{A}\varphi}$ NECA

Tabla 10.1: Axiomatización $\mathcal{L}_{\mathsf{Kc}_i,\mathsf{N}}$ de $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N}}$ con respecto a los NLTS^Us.

(KcE) Por TAUT, esto es equivalente a demostrar

$$\vdash (\mathsf{E}\psi \wedge \mathsf{Kc}_i(\psi, \varphi) \wedge \mathsf{A}\neg\varphi) \to \bot. \tag{10.1}$$

Para ello, estableceremos una serie de implicaciones y, utilizando la transitividad de la implicación, se tendrá la propiedad de arriba. Aplicando TAUT, se tiene que

$$\vdash (\mathsf{E}\psi \wedge \mathsf{Kc}_i(\psi, \varphi) \wedge \mathsf{A}\neg\varphi) \to (\mathsf{E}\psi \wedge \mathsf{Kc}_i(\psi, \varphi) \wedge \mathsf{A}\neg\varphi).$$

Tomando TAUT ($\vdash \neg\varphi \leftrightarrow (\varphi \to \bot)$ y $\vdash \psi \to \psi$) y NECA, en el lado derecho se sustituye $\mathsf{A}\neg\varphi$ por $\mathsf{A}(\varphi \to \bot)$ y se añade $\mathsf{A}(\psi \to \psi)$ que es una tautología. Por lo tanto,

$$\vdash (\mathsf{E}\psi \wedge \mathsf{Kc}_i(\psi, \varphi) \wedge \mathsf{A}\neg\varphi) \to (\mathsf{E}\psi \wedge \mathsf{A}(\psi \to \psi) \wedge \mathsf{Kc}_i(\psi, \varphi) \wedge \mathsf{A}(\varphi \to \bot)). \tag{10.2}$$

Por otro lado, usando una instancia de KcA, se tiene que

$$\vdash (\mathsf{A}(\psi \to \psi) \wedge \mathsf{Kc}_i(\psi, \varphi) \wedge \mathsf{A}(\varphi \to \bot)) \to (\mathsf{Kc}_i(\psi, \bot)).$$

Agregando $\mathsf{E}\psi$ a ambos lados (usando una instancia de TAUT),

$$\vdash (\mathsf{E}\psi \wedge \mathsf{A}(\psi \to \psi) \wedge \mathsf{Kc}_i(\psi, \varphi) \wedge \mathsf{A}(\varphi \to \bot)) \to (\mathsf{E}\psi \wedge \mathsf{Kc}_i(\psi, \bot)). \tag{10.3}$$

Por último, aplicando una instancia de KcN y agregando $\mathsf{E}\psi$ a ambos lados,

$$\vdash (\mathsf{E}\psi \wedge \mathsf{Kc}_i(\psi, \bot)) \to (\mathsf{E}\psi \wedge \mathsf{N}(\psi, \bot)).$$

Por definición de A y E,

$$\vdash (\mathsf{E}\psi \wedge \mathsf{Kc}_i(\psi, \bot)) \to (\neg\mathsf{A}\neg\psi \wedge \mathsf{A}\neg\psi).$$

Y por definición de $\bot$,

$$\vdash (\mathsf{E}\psi \wedge \mathsf{Kc}_i(\psi, \bot)) \to \bot. \tag{10.4}$$

Dadas las implicaciónes 10.2, 10.3 y 10.4, por transitividad se obtiene 10.1.

(NE) La demostración es similar a (KcE) sin usar KcN.

$\square$

Notar que $\mathsf{Kc}\bot$ es una instancia del axioma COND de la Proposición 5.1. Más aún, en [5] se lo considera como un axioma en vez de $\mathsf{NKc}\bot$. Sin embargo, al escribir esta tesis, se dio con que uno era derivable del otro y viceversa. Por lo tanto, la elección de uno de estos como axioma se tornaba indistinta. Para este trabajo, resultó más natural elegir $\mathsf{NKc}\bot$ ya que, como mencionamos, describe el vínculo entre la globalidad de A con el operador Kc_i. Con esto, damos las definiciones preliminares para demostrar completitud.

Definición 10.5. Sea $\mathbf{\Phi}$ el conjunto de todos los conjuntos maximales consistentes con $\mathcal{L}_{\mathsf{Kc}_i,\mathsf{N}}$. Para todo $\Gamma \in \mathbf{\Phi}$, se define:

$$\begin{aligned} \Gamma|_{\mathsf{N}} &:= \{\mathsf{N}(\psi,\varphi) \mid \mathsf{N}(\psi,\varphi) \in \Gamma\}, & \Gamma|_{\mathsf{A}} &:= \{\mathsf{A}\varphi \mid \mathsf{A}\varphi \in \Gamma\}, \\ \Gamma|_{\mathsf{Kc}_i} &:= \{\mathsf{Kc}_i(\psi,\varphi) \mid \mathsf{Kc}_i(\psi,\varphi) \in \Gamma\}, & \Gamma|_{\mathsf{Kc}} &:= \textstyle\bigcup_{i\in\mathsf{Agt}} \Gamma|_{\mathsf{Kc}_i}, \end{aligned}$$

$$\mathsf{Act}^\Gamma := \{\langle\psi,\varphi\rangle \mid \mathsf{N}(\psi,\varphi) \in \Gamma\}.$$

Notar que Act^Γ es numerable, por lo tanto es un conjunto de acciones adecuado para construir el modelo canónico como ocurría con Kh_i en el Capítulo 5. Con esto, definimos entonces el modelo canónico $\mathfrak{N}^\Gamma$ para un conjunto maximal consistente $\Gamma \in \mathbf{\Phi}$, tomando las ideas de los Capítulos 5 y 9.

Definición 10.6. Sea $\Gamma \in \mathbf{\Phi}$ un conjunto maximal consistente fórmulas, el *modelo canónico* para Γ es una tupla $\mathfrak{N}^\Gamma = \langle \mathrm{S}^\Gamma, \mathrm{R}^\Gamma, \mathrm{V}^\Gamma, \mathrm{U}^\Gamma, \mathrm{N}^\Gamma\rangle$ sobre Act^Γ, Agt y Prop donde:

- $\mathrm{S}^\Gamma = \{\Delta \in \mathbf{\Phi} \mid \Delta|_{\mathsf{A}} = \Gamma|_{\mathsf{A}}\}$;
- $\mathrm{R}^\Gamma_{\langle\psi,\varphi\rangle} = \{(\Delta_1,\Delta_2) \in \mathrm{S}^\Gamma \times \mathrm{S}^\Gamma \mid \mathsf{N}(\psi,\varphi) \in \Gamma, \psi \in \Delta_1, \varphi \in \Delta_2\}$;
- $\mathrm{R}^\Gamma = \{\mathrm{R}^\Gamma_{\langle\psi,\varphi\rangle} \mid \mathsf{N}(\psi,\varphi) \in \Gamma\}$;
- $\mathrm{V}^\Gamma(\Delta) = \{p \in \mathsf{Prop} \mid p \in \Delta\}$;
- $\mathrm{U}^\Gamma(i) = \{\{\langle\psi,\varphi\rangle\} \mid \mathsf{Kc}_i(\psi,\varphi) \in \Gamma\} \cup \{\emptyset\}$; y
- $\mathrm{N}^\Gamma = \{\langle\psi,\varphi\rangle \mid \mathsf{N}(\psi,\varphi) \in \Gamma\}$.

Notar que el modelo $\mathfrak{N}^\Gamma$ está generado por fórmulas de la forma $\mathsf{A}\varphi$, es decir, las fórmulas globales de Γ. Esto se debe a que ahora, a diferencia de lo ocurrido con $\mathsf{L}_{\mathsf{Kh}_i}$, contamos con más de un operador que genera nuestro modelo, similar a lo que sucedía con $\mathsf{L}_{\mathsf{Kh}_i,\square}$ cuando agregábamos las modalidades $[a]$.

Proposición 10.3. *Se tienen las siguientes propiedades para* $\mathfrak{N}^\Gamma$*:*

1. $\{\langle\psi,\varphi\rangle\} \in \mathrm{U}^\Gamma(i)$ *implica* $\langle\psi,\varphi\rangle \in \mathrm{N}^\Gamma$*;*
2. $\langle\psi,\varphi\rangle \in \mathrm{N}^\Gamma$ *sii* $\langle\psi,\varphi\rangle \in \mathsf{Act}^\Gamma$*; y*
3. $\mathrm{R}^\Gamma_{\langle\psi,\varphi\rangle} \neq \emptyset$ *implica* $\mathsf{N}(\psi,\varphi) \in \Gamma$.

Demostración. El primer item se cumple por el axioma KcN. El segundo y el tercero son por la definición de $\mathfrak{N}^\Gamma$. $\square$

Con estas propiedades, se demuestra que $\mathfrak{N}^\Gamma$ es un NLTS^U.

Proposición 10.4. $\mathfrak{N}^\Gamma = \langle \mathrm{S}^\Gamma, \mathrm{R}^\Gamma, \mathrm{V}^\Gamma, \mathrm{U}^\Gamma, \mathrm{N}^\Gamma\rangle$ *es un* NLTS^U.

Demostración. Claramente, $\langle \mathrm{S}^\Gamma, \mathrm{R}^\Gamma, \mathrm{V}^\Gamma\rangle$ es un LTS ($\mathrm{S}^\Gamma \neq \emptyset$, dado que $\Gamma \in \mathrm{S}^\Gamma$). Para probar que $\langle \mathrm{S}^\Gamma, \mathrm{R}^\Gamma, \mathrm{V}^\Gamma, \mathrm{N}^\Gamma\rangle$ es un NLTS^U (Definición 10.2), basta con ver que U^Γ y N^Γ cumplen las condiciones.

- Por definición, $\emptyset \in \mathrm{U}^\Gamma(i)$. Dados $\pi_1, \pi_2 \in \mathrm{U}^\Gamma_i - \{\emptyset\}$ tales que $\pi_1 = \{\langle\psi_1,\varphi_1\rangle\} \neq \{\langle\psi_2,\varphi_2\rangle\} = \pi_2$, como π_1, π_2 son conjuntos unitarios, se tiene que $\pi_1 \cap \pi_2 = \emptyset$.

- Por DN, $\mathsf{N}(\varphi, \top) \in \Gamma$ para toda fórmula φ. Particularmente, $\mathsf{N}(\top, \top) \in \Gamma$. Por definición, $\mathrm{R}^{\Gamma}_{\langle\top,\top\rangle} = \mathrm{S}^{\Gamma} \times \mathrm{S}^{\Gamma}$ y $\langle\top, \top\rangle \in \mathrm{N}^{\Gamma}$. Con esto, $[\![\top]\!]^{\mathfrak{N}^{\Gamma}} = \mathrm{S}^{\Gamma} = \mathrm{SE}(\langle\top, \top\rangle)$ y $\langle\top, \top\rangle$ es fuertemente ejecutable en S.

Por lo tanto, $\mathfrak{N}^{\Gamma}$ es un NLTS^{U}. $\square$

Las siguientes son algunas propiedades estándar de $\mathfrak{N}^{\Gamma}$.

Proposición 10.5. *Sean* $\Delta_1, \Delta_2 \in \mathrm{S}^{\Gamma}$ *y* $\mathsf{X} \in \{\mathsf{Kc}_i, \mathsf{N}\}$*; se tiene que* $\Delta_1|_{\mathsf{X}} = \Delta_2|_{\mathsf{X}} = \Gamma|_{\mathsf{X}}$.

Demostración. Veamos para el caso $\mathsf{X} = \mathsf{Kc}_i$. Sean $\Delta_1, \Delta_2 \in \mathrm{S}^{\Gamma}$, veamos que $(\Delta_1)|_{\mathsf{Kc}} = (\Delta_2)|_{\mathsf{Kc}}$.

($\subseteq$) Supongamos que $\mathsf{Kc}_i(\psi, \varphi) \in (\Delta_1)|_{\mathsf{Kc}_i}$. Como Δ_1 es un conjunto maximal consistente, por $\mathsf{4KcA}$ se tiene que $\mathsf{AKc}_i(\psi, \varphi) \in (\Delta_1)|_{\mathsf{A}}$. Por lo tanto, por la Definición 10.6, se tiene que $\mathsf{AKc}_i(\psi, \varphi) \in \Gamma|_{\mathsf{A}}$, y con esto $\mathsf{AKc}_i(\psi, \varphi) \in (\Delta_2)|_{\mathsf{A}}$. Por TA, se tiene que $\mathsf{Kc}_i(\psi, \varphi) \in (\Delta_2)|_{\mathsf{Kc}_i}$.

($\supseteq$) Para la otra inclusión se muestra por contrarrecíproca. Supongamos que $\mathsf{Kc}_i(\psi, \varphi) \notin (\Delta_1)|_{\mathsf{Kc}_i}$. Como Δ_1 es un conjunto maximal consistente, $\neg\mathsf{Kc}_i(\psi, \varphi) \in \Delta_1$. Por $\mathsf{5KcA}$, se tiene que $\mathsf{A}\neg\mathsf{Kc}_i(\psi, \varphi) \in (\Delta_1)|_{\mathsf{A}}$, y por la Definición 10.6, $\mathsf{A}\neg\mathsf{Kc}_i(\psi, \varphi) \in \Gamma|_{\mathsf{A}}$. Con esto, $\mathsf{A}\neg\mathsf{Kc}_i(\psi, \varphi) \in (\Delta_2)|_{\mathsf{A}}$, y por TA, $\neg\mathsf{Kc}_i(\psi, \varphi) \in \Delta_2$. Por lo tanto, $\mathsf{Kc}_i(\psi, \varphi)(\Delta_2)|_{\mathsf{Kc}_i}$.

El caso $\mathsf{X} = \mathsf{N}$ es similar usando $\mathsf{4NA}$ y $\mathsf{5NA}$. $\square$

La Proposición 10.5 habla sobre las relaciones globales entre los estados de $\mathfrak{N}^{\Gamma}$. Con esto, demostramos algunas propiedades sobre a estructura de $\mathfrak{N}^{\Gamma}$.

Proposición 10.6. *Si* $\mathrm{R}^{\Gamma}_{\langle\psi,\varphi\rangle}(\Delta) \neq \emptyset$*, entonces, para todo* $\Delta' \in \mathrm{S}^{\Gamma}$, $\varphi \in \Delta'$ *implica* $\Delta' \in \mathrm{R}^{\Gamma}_{\langle\psi,\varphi\rangle}(\Delta)$.

Demostración. Si Δ tiene un sucesor vía $\mathrm{R}^{\Gamma}_{\langle\psi,\varphi\rangle}$, entonces $\psi \in \Delta$ y $\mathsf{N}(\psi, \varphi) \in \Gamma$. Entonces, para todo $\Delta' \in \mathrm{S}^{\Gamma}$ con $\varphi \in \Delta'$, se tiene que $(\Delta, \Delta') \in \mathrm{R}^{\Gamma}_{\langle\psi,\varphi\rangle}$. $\square$

Proposición 10.7. *Sea* φ *una fórmula de* $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N}}$. *Si* $\varphi \in \Delta$ *para todo* $\Delta \in \mathrm{S}^{\Gamma}$*, entonces* $\mathsf{A}\varphi \in \Delta$ *para todo* $\Delta \in \mathrm{S}^{\Gamma}$.

Demostración. La prueba es análoga a la Proposición 9.5, extendiendo la Definición 10.5 con $\Delta|_{\neg\mathsf{A}} = \{\neg\mathsf{A}\varphi \mid \neg\mathsf{A}\varphi \in \Delta\}$ y usando $\mathsf{4NA}$ y $\mathsf{5NA}$. $\square$

Proposición 10.8. *Sean* ψ, ψ', φ' *fórmulas de* $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N}}$. *Si para todo* $\Delta \in \mathrm{S}^{\Gamma}$ *tal que* $\psi \in \Delta$*, tiene un sucesor vía* $\mathrm{R}^{\Gamma}_{\langle\psi',\varphi'\rangle}$. *Entonces* $\mathsf{A}(\psi \to \psi') \in \Delta$ *para todo* $\Delta \in \mathrm{S}^{\Gamma}$.

Demostración. Análoga a la Proposición 5.6 usando la Proposición 10.7. $\square$

Proposición 10.9. *Sea* $\mathsf{X} \in \{\mathsf{Kc}_i, \mathsf{N}\}$*, si existe* $\Delta \in \mathrm{S}^{\Gamma}$ *tal que* $\{\psi, \mathsf{X}(\psi, \varphi)\} \subset \Delta$*, entonces existe* $\Delta' \in \mathrm{S}^{\Gamma}$ *tal que* $\varphi \in \Delta'$.

Demostración. Sea $\Delta \in \mathrm{S}^{\Gamma}$ tal que $\{\psi, \mathsf{X}(\psi, \varphi)\} \subset \Delta$. Por la contrarrecíproca de TA, se tiene que $\mathsf{E}\psi \in \Delta$. Supongamos que no existe $\Delta' \in \mathrm{S}^{\Gamma}$ tal que $\varphi \in \Delta'$. Luego, por consistencia maximal, para todo $\Delta' \in \mathrm{S}^{\Gamma}$, $\neg\varphi \in \Delta'$. Por la Proposición 10.7, para todo $\Delta' \in \mathrm{S}^{\Gamma}$, $\mathsf{A}\neg\varphi \in \Delta'$ y $\mathsf{A}\neg\varphi \in \Delta$. Como $\{\mathsf{E}\psi, \mathsf{X}(\psi, \varphi), \mathsf{A}\neg\varphi\} \subset \Delta$, por la Proposición 10.2 y consistencia maximal, $\bot \in \Delta$. Pero es contradictorio pues Δ es un conjunto consistente. Por lo tanto, existe $\Delta' \in \mathrm{S}^{\Gamma}$ tal que $\varphi \in \Delta'$. $\square$

Con estas propiedades, procedemos al "truth lemma" para $\mathfrak{N}^{\Gamma}$.

Lema 10.1. *Sean* $\Gamma \in \mathbf{\Phi}$ *y* $\mathfrak{N}^{\Gamma} = \langle \mathrm{S}^{\Gamma}, \mathrm{R}^{\Gamma}, \mathrm{V}^{\Gamma}, \mathrm{U}^{\Gamma}, \mathrm{N}^{\Gamma}\rangle$. *Para todo* $\Theta \in \mathrm{S}^{\Gamma}$ *y* φ *fórmula de* $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N}}$,

$$\mathfrak{N}^{\Gamma}, \Theta \models \varphi \qquad \textit{si y sólo si} \qquad \varphi \in \Theta.$$

Demostración. Sea $\mathfrak{N}^{\Gamma} = \langle \mathrm{S}^{\Gamma}, \mathrm{R}^{\Gamma}, \mathrm{V}^{\Gamma}, \mathrm{U}^{\Gamma}, \mathrm{N}^{\Gamma}\rangle$. La prueba es por inducción en φ. Los casos atómicos y para los operadores booleanos son directos, por lo que sólo se verá el caso con el operador Kc_i (el caso para N es similar).

- Caso $\mathsf{Kc}_i(\psi, \varphi)$: ($\Rightarrow$) Supongamos que $\mathfrak{N}^{\Gamma}, \Theta \models \mathsf{Kc}_i(\psi, \varphi)$. Luego existe $\pi \in \mathrm{U}^{\Gamma}(i)$ tal que: $\pi \subseteq \mathrm{N}^{\Gamma}$, $[\![\psi]\!]^{\mathfrak{N}^{\Gamma}} \subseteq \mathrm{SE}(\pi)$ y $\mathrm{R}_{\pi}([\![\psi]\!]^{\mathfrak{N}^{\Gamma}}) \subseteq [\![\varphi]\!]^{\mathfrak{N}^{\Gamma}}$. Consideramos dos situaciones:

- $[\![\psi]\!]^{\mathfrak{N}^\Gamma} = \emptyset$: Por hipótesis inductiva, $\neg\psi \in \Delta$ para todo $\Delta \in \mathrm{S}^\Gamma$. Por la Proposición 10.7, $\mathsf{A}\neg\psi \in \Delta$ para todo $\Delta \in \mathrm{S}^\Gamma$, y con esto $\mathsf{A}(\psi \to \bot) \in \Delta$. Como $\bot \to \varphi$ es una tautología, por NECA, $\mathsf{A}(\bot \to \varphi) \in \Delta$. Tomando una instancia de KcA, $(\mathsf{A}(\psi \to \bot) \land \mathsf{Kc}_i(\bot, \bot) \land \mathsf{A}(\bot \to \varphi)) \to \mathsf{Kc}_i(\psi, \varphi) \in \Delta$. Luego, por $\mathsf{Kc}\bot$ y MP, se tiene que $\mathsf{Kc}_i(\psi, \varphi) \in \Delta$ para todo $\Delta \in \mathrm{S}^\Gamma$. Por lo tanto, $\mathsf{Kc}_i(\psi, \varphi) \in \Theta$.
- $[\![\psi]\!]^{\mathfrak{N}^\Gamma} \neq \emptyset$: Luego, $\pi = \{\langle \psi', \varphi' \rangle\}$, dado que $\pi = \emptyset$ forzaría a que $[\![\psi]\!]^{\mathfrak{N}^\Gamma} = \emptyset$. Por definición de U^Γ, $\mathsf{Kc}_i(\psi', \varphi') \in \Gamma$, y usando KcN, $\mathrm{R}^\Gamma_{\langle \psi', \varphi' \rangle}$ está definida. Sea $\Delta \in \mathrm{S}^\Gamma$ tal que $\mathfrak{N}^\Gamma, \Delta \models \psi$. Por hipótesis inductiva, $\psi \in \Delta$. Más aún, para todo $\Delta \in \mathrm{S}^\Gamma$, si $\psi \in \Delta$ entonces:
  - Δ tiene un sucesor vía $\mathrm{R}^\Gamma_{\langle \psi', \varphi' \rangle}$, y
  - para todo $\Delta' \in \mathrm{S}^\Gamma$ tal que $(\Delta, \Delta') \in \mathrm{R}^\Gamma_{\langle \psi', \varphi' \rangle}$, $\varphi \in \Delta'$ (hipótesis inductiva).

  Con esto deducimos tres partes:
  1. Como para todo $\Delta \in \mathrm{S}^\Gamma$ tal que $\psi \in \Delta$, tiene un sucesor vía $\mathrm{R}^\Gamma_{\langle \psi', \varphi' \rangle}$, por la Proposición 10.8, $\mathsf{A}(\psi \to \psi') \in \Delta$ para todo $\Delta \in \mathrm{S}^\Gamma$.
  2. Por la Proposición 10.6, cada $\Delta' \in \mathrm{S}^\Gamma$ tal que $\varphi' \in \Delta'$ puede ser alcanzado vía $\mathrm{R}^\Gamma_{\langle \psi', \varphi' \rangle}$ desde un $\Delta \in \mathrm{S}^\Gamma$ tal que $\psi \in \Delta$ (pues, $[\![\psi]\!]^{\mathfrak{N}^\Gamma} \neq \emptyset$) y por lo tanto $\varphi \in \Delta'$. Por consistencia maximal, para todo $\Delta' \in \mathrm{S}^\Gamma$, $\varphi' \to \varphi \in \Delta'$. Usando la Proposición 10.7, para todo $\Delta' \in \mathrm{S}^\Gamma$, $\mathsf{A}(\varphi' \to \varphi) \in \Delta'$.
  3. Dado $\Delta \in \mathrm{S}^\Gamma$, como $\Delta|_{\mathsf{Kc}_i} = \Gamma|_{\mathsf{Kc}_i}$ y $\mathsf{Kc}_i(\psi', \varphi') \in \Gamma$, entonces $\mathsf{Kc}_i(\psi', \varphi') \in \Delta$.

  Finalmente, para todo $\Delta \in \mathrm{S}^\Gamma$, $\{\mathsf{A}(\psi \to \psi'), \mathsf{Kc}_i(\psi', \varphi'), \mathsf{A}(\varphi' \to \varphi)\} \subset \Delta$. Usando KcA y MP, se tiene que $\mathsf{Kc}_i(\psi, \varphi) \in \Theta$.

($\Leftarrow$) Supongamos que $\mathsf{Kc}_i(\psi, \varphi) \in \Theta$. Luego, por la Proposición 10.5, $\mathsf{Kc}_i(\psi, \varphi) \in \Delta$ para todo $\Delta \in \mathrm{S}^\Gamma$. Más aún, $\mathsf{Kc}_i(\psi, \varphi) \in \Gamma$ y $\mathrm{R}^\Gamma_{\langle \psi, \varphi \rangle}$ está definida. Para probar que $\mathfrak{N}^\Gamma, \Theta \models \mathsf{Kc}_i(\psi, \varphi)$, consideramos dos casos:

- No existe Δ tal que $\psi \in \Delta$: Por hipótesis inductiva, $[\![\psi]\!]^{\mathfrak{N}^\Gamma} = \emptyset$. Usando $\pi = \emptyset$, se tiene trivialmente que $\mathfrak{N}^\Gamma, \Theta \models \mathsf{Kc}_i(\psi, \varphi)$.
- Existe Δ tal que $\psi \in \Delta$: Por la Proposición 10.9, existe Δ' tal que $\varphi \in \Delta'$. Por hipótesis inductiva, $\mathfrak{N}^\Gamma, \Delta \models \psi$ y $\mathfrak{N}^\Gamma, \Delta' \models \varphi$. Como está definido, $\pi = \{\langle \psi, \varphi \rangle\} \in \mathrm{U}^\Gamma$ es fuertemente ejecutable en todos los estados donde se cumple ψ (dado que existe un sucesor Δ' vía $\mathrm{R}^\Gamma_{\langle \psi, \varphi \rangle}$), alcanza desde estos sólo estados donde se cumple φ vía π (por construcción de $\mathrm{R}^\Gamma_{\langle \psi, \varphi \rangle}$), y $\{\langle \psi, \varphi \rangle\} \subseteq \mathrm{N}^\Gamma$ (Proposición 10.3). Por lo tanto, $\mathfrak{N}^\Gamma, \Theta \models \mathsf{Kc}_i(\psi, \varphi)$.

□

Con el Lema 10.1, finalmente se obtienen las propiedades buscadas de esta axiomatización.

Teorema 10.1. *El sistema axiomático $\mathcal{L}_{\mathsf{Kc}_i,\mathsf{N}}$ (Tabla 10.1) es correcto y fuertemente completo para $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N}}$ con respecto a la clase de todos los* NLTS^U*s.*

Demostración. Para la correctitud, basta con mostrar que los axiomas del sistema son validos y que sus reglas preservan la validez, que resulta ser directo. Para la completitud fuerte, sea Γ' un conjunto de fórmulas consistente con $\mathcal{L}_{\mathsf{Kc}_i,\mathsf{N}}$. Teniendo en cuenta la Observación 5.2 aplicada a $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N}}$, Γ' puede ser extendido a un conjunto maximal consistente con $\mathcal{L}_{\mathsf{Kc}_i,\mathsf{N}}$ $\Gamma \supseteq \Gamma'$. Por el Lema 10.1, $\mathfrak{N}^\Gamma, \Gamma \models \Gamma'$ y Γ' es satisfacible. Dado que $\mathfrak{N}^\Gamma$ es un NLTS^U (Proposición 10.4), esto termina de demostrar la completitud fuerte. □

10.4. Complejidad

En esta sección se analiza la complejidad computacional del problema de satisfacibilidad de $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N}}$ bajo la semántica basada en NLTS^U, tal como lo realizamos en [5]. Similarmente a la Sección 7.2, se establecerá que dicho problema es NP-completo mostrando que, si una fórmula es satisfacible, entonces existe un modelo de tamaño polinomial que la satisface. Además, introduciremos un algoritmo de model checking que corre en tiempo polinomial.

Dada una fórmula, se mostrará que es posible seleccionar un fragmento del modelo canónico relevante para su evaluación. Este modelo seleccionado bastará para preservar la satisfacibilidad de la fórmula. Es más, el tamaño de este modelo será polinomial al tamaño de la fórmula dada.

Definición 10.7 (Función de selección)**.** Sea $\mathfrak{N}^\Gamma = \langle \mathrm{S}^\Gamma, \mathrm{R}^\Gamma, \mathrm{V}^\Gamma, \mathrm{U}^\Gamma, \mathrm{N}^\Gamma \rangle$ un modelo canónico para un conjunto maximal consistente Γ (Definición 10.6). Sean $w \in \mathrm{S}^\Gamma$ y una fórmula φ de $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N}}$. Se define $\mathsf{Act}_\varphi = \{\langle \theta_1, \theta_2 \rangle \in \mathsf{Act}^\Gamma \mid \mathsf{X}(\theta_1, \theta_2)$ es subfórmula de $\varphi\} \cup \{\langle \top, \top \rangle\}$, con $\mathsf{X} \in \{\mathsf{Kc}_i, \mathsf{N}\}$. Una *función de selección canónica* sel_w^φ es una función que toma un modelo $\mathfrak{N}^\Gamma$, $w \in \mathrm{S}^\Gamma$ y una fórmula φ, y devuelve un conjunto de estados $\mathrm{S}' \subseteq \mathrm{S}^\Gamma$ de la siguiente manera:

1. $\mathsf{sel}_w^\varphi(p) = \{w\}$;
2. $\mathsf{sel}_w^\varphi(\neg\varphi_1) = \mathsf{sel}_w^\varphi(\varphi_1)$
3. $\mathsf{sel}_w^\varphi(\varphi_1 \vee \varphi_2) = \mathsf{sel}_w^\varphi(\varphi_1) \cup \mathsf{sel}_w^\varphi(\varphi_2)$;
4. Si $[\![\mathsf{X}(\varphi_1, \varphi_2)]\!]^{\mathfrak{N}^\Gamma} \neq \emptyset$ y $[\![\varphi_1]\!]^{\mathfrak{N}^\Gamma} = \emptyset$ para $\mathsf{X} \in \{\mathsf{Kc}_i, \mathsf{N}\}$:
   $\mathsf{sel}_w^\varphi(\mathsf{X}(\varphi_1, \varphi_2)) = \{w\}$;
5. Si $[\![\mathsf{X}(\varphi_1, \varphi_2)]\!]^{\mathfrak{N}^\Gamma} \neq \emptyset$ y $[\![\varphi_1]\!]^{\mathfrak{N}^\Gamma} \neq \emptyset$ para $\mathsf{X} \in \{\mathsf{Kc}_i, \mathsf{N}\}$:
   $\mathsf{sel}_w^\varphi(\mathsf{X}(\varphi_1, \varphi_2)) = \{w_1, w_2\} \cup \mathsf{sel}_{w_1}^\varphi(\varphi_1) \cup \mathsf{sel}_{w_2}^\varphi(\varphi_2)$,
   donde w_1, w_2 es tal que $(w_1, w_2) \in \mathrm{R}^\Gamma_{\langle \varphi_1, \varphi_2 \rangle}$;
6. Si $[\![\mathsf{Kc}_i(\varphi_1, \varphi_2)]\!]^{\mathfrak{N}^\Gamma} = \emptyset$ (notar que $[\![\varphi_1]\!]^{\mathfrak{N}^\Gamma} \neq \emptyset$):
   Para cada $\{\langle \psi_1, \psi_2 \rangle\} = \{a\} \in \mathrm{U}^\Gamma(i)$ tal que $a \in \mathsf{Act}_\varphi$:
   a) si $[\![\varphi_1]\!]^{\mathfrak{N}^\Gamma} \not\subseteq \mathrm{SE}(a)$,
   se agrega $\mathsf{sel}_{w_1}^\varphi(\varphi_1) \cup \{w_1\}$ a $\mathsf{sel}_w^\varphi(\mathsf{Kc}_i(\varphi_1, \varphi_2))$,
   donde $w_1 \in [\![\varphi_1]\!]^{\mathfrak{N}^\Gamma}$ y $w_1 \notin \mathrm{SE}(a)$;
   b) si $[\![\varphi_1]\!]^{\mathfrak{N}^\Gamma} \subseteq \mathrm{SE}(a)$ pero $\mathrm{R}^\Gamma_a([\![\varphi_1]\!]^{\mathfrak{N}^\Gamma}) \not\subseteq [\![\varphi_2]\!]^{\mathfrak{N}^\Gamma}$,
   se agrega $\{w_1, w_2\} \cup \mathsf{sel}_{w_1}^\varphi(\varphi_1) \cup \mathsf{sel}_{w_2}^\varphi(\varphi_2)$ a $\mathsf{sel}_w^\varphi(\mathsf{Kc}_i(\varphi_1, \varphi_2))$,
   donde $w_1 \in [\![\varphi_1]\!]^{\mathfrak{N}^\Gamma}$, $w_2 \in \mathrm{R}^\Gamma_a(w_1)$ y $w_2 \notin [\![\varphi_2]\!]^{\mathfrak{N}^\Gamma}$;
7. Si $[\![\mathsf{N}(\varphi_1, \varphi_2)]\!]^{\mathfrak{N}^\Gamma} = \emptyset$ (notar que $[\![\varphi_1]\!]^{\mathfrak{N}^\Gamma} \neq \emptyset$):
   Para cada $\langle \psi_1, \psi_2 \rangle = a \in \mathrm{N}^\Gamma \cap \mathsf{Act}_\varphi$:
   a) si $[\![\varphi_1]\!]^{\mathfrak{N}^\Gamma} \not\subseteq \mathrm{SE}(a)$,
   se agrega $\{w_1\} \cup \mathsf{sel}_{w_1}^\varphi(\varphi_1)$ a $\mathsf{sel}_w^\varphi(\mathsf{N}(\varphi_1, \varphi_2))$,
   donde $w_1 \in [\![\varphi_1]\!]^{\mathfrak{N}^\Gamma}$ y $w_1 \notin \mathrm{SE}(a)$;
   b) si $[\![\varphi_1]\!]^{\mathfrak{N}^\Gamma} \subseteq \mathrm{SE}(a)$ pero $\mathrm{R}^\Gamma_a([\![\varphi_1]\!]^{\mathfrak{N}^\Gamma}) \not\subseteq [\![\varphi_2]\!]^{\mathfrak{N}^\Gamma}$,
   se agrega $\{w_1, w_2\} \cup \mathsf{sel}_{w_1}^\varphi(\varphi_1) \cup \mathsf{sel}_{w_2}^\varphi(\varphi_2)$ a $\mathsf{sel}_w^\varphi(\mathsf{N}(\varphi_1, \varphi_2))$,
   donde $w_1 \in [\![\varphi_1]\!]^{\mathfrak{N}^\Gamma}$, $w_2 \in \mathrm{R}^\Gamma_a(w_1)$ y $w_2 \notin [\![\varphi_2]\!]^{\mathfrak{N}^\Gamma}$.

Esta función de selección elige los estados suficientes del modelo canónico de tal manera que preserva el valor de verdad de las subfórmulas de una fórmula dada. Para dar con esta propiedad, será necesario seleccionar un conjunto adecuado de planes normativos y los conjuntos de incertidumbre. Esto se lo detalla en la siguiente definición.

Definición 10.8 (Modelo seleccionado)**.** Sea $\mathfrak{N}^\Gamma$ un modelo canónico para un conjunto maximal consistente Γ, w un estado de $\mathfrak{N}^\Gamma$, y φ una fórmula en $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N}}$. Sea sel_w^φ una función de selección, se define el *modelo seleccionado por* sel_w^φ como $\mathfrak{N}_w^\varphi = \langle \mathrm{S}_w^\varphi, \mathrm{R}_w^\varphi, \mathrm{V}_w^\varphi, \mathrm{U}_w^\varphi, \mathrm{N}_w^\varphi \rangle$, donde

- $\mathrm{S}_w^\varphi = \mathsf{sel}_w^\varphi(\varphi)$;
- $(\mathrm{R}_w^\varphi)_{\langle \theta_1, \theta_2 \rangle} = \mathrm{R}^\Gamma_{\langle \theta_1, \theta_2 \rangle} \cap (\mathrm{S}_w^\varphi \times \mathrm{S}_w^\varphi)$, para cada $\langle \theta_1, \theta_2 \rangle \in \mathsf{Act}_\varphi$;
- $\mathrm{N}_w^\varphi = \mathrm{N}^\Gamma \cap \mathsf{Act}_\varphi$;

- $(\mathrm{U}_w^\varphi)(i) = \{\{a\} \in \mathrm{U}^\Gamma(i) \mid a \in \mathsf{Act}_\varphi\} \cup \{\emptyset\}$, para $i \in \mathsf{Agt}$;
- V_w^φ es la restricción de V^Γ a S_w^φ.

Análogamente a la Proposición 7.1, $\mathfrak{N}_w^\varphi$ la estructura definida es un NLTS^U.

Proposición 10.10. *$\mathfrak{N}_w^\varphi = \langle \mathrm{S}_w^\varphi, \mathrm{R}_w^\varphi, \mathrm{V}_w^\varphi, \mathrm{U}_w^\varphi, \mathrm{N}_w^\varphi \rangle$ es un NLTS^U. Más aún, $\pi \in \mathrm{U}_w^\varphi(i)$ implica que $\pi \subseteq \mathrm{N}_w^\varphi$.*

Demostración. La estructura $\langle \mathrm{S}_w^\varphi, \mathrm{R}_w^\varphi, \mathrm{V}_w^\varphi \rangle$ es un LTS dado que $\mathsf{sel}_w^\varphi(\varphi) \neq \emptyset$. Como preservamos $\langle \top, \top \rangle$ en N_w^φ y (U_w^φ) es una restricción de de U^Γ a Act_φ, es fácil ver que se cumplen las condiciones de la Definición 10.2. Por lo tanto, $\mathfrak{N}_w^\varphi$ es un NLTS^U.

Sea $\pi \in \mathrm{U}_w^\varphi(i)$. Si $\pi = \emptyset$, es fácil ver que $\pi \subseteq \mathrm{N}_w^\varphi$. Si en cambio $\pi = \{\langle \theta_1, \theta_2 \rangle\}$, entonces $\{\langle \theta_1, \theta_2 \rangle\} \in \mathrm{U}^\Gamma(i)$ y $\langle \theta_1, \theta_2 \rangle \in \mathsf{Act}_\varphi$. Independientemente de si $\mathsf{Kc}_i(\theta_1, \theta_2)$ ó $\mathsf{N}(\theta_1, \theta_2)$ es subfórmula de φ, por el primer ítem de la Proposición 10.3, $\langle \theta_1, \theta_2 \rangle \in \mathrm{N}^\Gamma$. Con esto, $\langle \theta_1, \theta_2 \rangle \in \mathrm{N}_w^\varphi$ y se tiene que $\pi \subseteq \mathrm{N}_w^\varphi$. $\square$

Con esto, podemos demostrar que el modelo seleccionado preserva la satisfacibilidad de la fórmula del modelo canónico.

Proposición 10.11. *Sea $\mathfrak{N}^\Gamma$ un modelo canónico, w un estado de $\mathfrak{N}^\Gamma$ y φ una fórmula de $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N}}$. Sea $\mathfrak{N}_w^\varphi$ un modelo seleccionado por una función de selección sel_w^φ. Entonces, $\mathfrak{N}^\Gamma, w \models \varphi$ implica que para toda ψ subfórmula de φ, y para toda $v \in \mathrm{S}_w^\varphi$ se tiene que $\mathfrak{N}^\Gamma, v \models \psi$ si y sólo si $\mathfrak{N}_w^\varphi, v \models \psi$. Más aún, $\mathfrak{N}_w^\varphi$ es de tamaño polinomial al tamaño de la fórmula φ.*

Demostración. La demostración se hace por inducción en el tamaño de la fórmula. φ. Los casos atómicos y demás operadores booleanos son estándar y análogos a los vistos en la Proposición 7.2. Para el caso de Kc_i, no es necesario comprobar que $\pi \subseteq \mathrm{N}_w^\varphi$ por la Proposición 10.10:

- Caso $\psi = \mathsf{Kc}_i(\psi_1, \psi_2)$: Si $\mathfrak{N}^\Gamma, v \models \mathsf{Kc}_i(\psi_1, \psi_2)$, hay dos posibilidades:
  - $[\![\psi_1]\!]^{\mathfrak{N}^\Gamma} = \emptyset$: Por hipótesis inductiva, $[\![\psi_1]\!]^{\mathfrak{N}_w^\varphi} = \emptyset$ y basta con elegir un $\pi' = \emptyset$ como testigo normativo de i.
  - $[\![\psi_1]\!]^{\mathfrak{N}^\Gamma} \neq \emptyset$: Por el Lema 10.1 y la definición de Act_φ, se tiene que $\{\langle \psi_1, \psi_2 \rangle\}$ es un testigo normativo de i para $\mathsf{Kc}_i(\psi_1, \psi_2)$ en $\mathfrak{N}^\Gamma$ y $\langle \psi_1, \psi_2 \rangle \in \mathsf{Act}_\varphi$ (dado que es una subfórmula de φ). Como $[\![\psi_1]\!]^{\mathfrak{N}^\Gamma} \neq \emptyset$, existen $w_1, w_2 \in \mathrm{S}^\Gamma$ tales que $(w_1, w_2) \in \mathrm{R}^\Gamma_{\langle \psi_1, \psi_2 \rangle}$.

    Por definición de $\mathfrak{N}_w^\varphi$, $\{\langle \psi_1, \psi_2 \rangle\} \in (\mathrm{U}_w^\varphi)(i)$ y $(\mathrm{R}_w^\varphi)_{\langle \psi_1, \psi_2 \rangle}$ está definida. Además, por definición de sel_w^φ, Ítem 5, existen $w_1', w_2' \in \mathrm{S}_w^\varphi$ tales que $(w_1', w_2') \in (\mathrm{R}_w^\varphi)_{\langle \psi_1, \psi_2 \rangle}$. Sea $v_1 \in [\![\psi_1]\!]^{\mathfrak{N}_w^\varphi} \subseteq [\![\psi_1]\!]^{\mathfrak{N}^\Gamma}$. Entonces, $v_1 \in \mathrm{SE}^{\mathfrak{N}^\Gamma}(\langle \psi_1, \psi_2 \rangle)$ y $\mathrm{R}^\Gamma_{\langle \psi_1, \psi_2 \rangle}(v_1) \subseteq [\![\psi_2]\!]^{\mathfrak{N}^\Gamma}$. Para todo $v_2 \in \mathrm{R}^\Gamma_{\langle \psi_1, \psi_2 \rangle}(v_1)$ se tiene que $v_2 \in [\![\psi_2]\!]^{\mathfrak{N}^\Gamma}$, particularmente $v_2 = w_2'$. Con esto, $w_2' \in (\mathrm{R}_w^\varphi)_{\langle \psi_1, \psi_2 \rangle}(v_1)$ y $v_1 \in \mathrm{SE}^{\mathfrak{N}_w^\varphi}(\{\langle \psi_1, \psi_2 \rangle\})$.

    Supongamos que $(\mathrm{R}_w^\varphi)_{\langle \psi_1, \psi_2 \rangle}(v_1) = \mathrm{R}^\Gamma_{\langle \psi_1, \psi_2 \rangle}(v_1) \cap \mathrm{S}_w^\varphi \not\subseteq [\![\psi_2]\!]^{\mathfrak{N}_w^\varphi}$. Sea $v_2 \in (\mathrm{R}_w^\varphi)_{\langle \psi_1, \psi_2 \rangle}(v_1)$ tal que $v_2 \notin [\![\psi_2]\!]^{\mathfrak{N}_w^\varphi}$, se tiene que $(\mathrm{R}_w^\varphi)_{\langle \psi_1, \psi_2 \rangle}(v_1) \subseteq \mathrm{R}^\Gamma_{\langle \psi_1, \psi_2 \rangle}(v_1)$, y por hipótesis inductiva $v_2 \notin [\![\psi_2]\!]^{\mathfrak{N}^\Gamma}$. Con esto, $\{\langle \psi_1, \psi_2 \rangle\}$ no es un testigo válido en $\mathfrak{N}^\Gamma$, lo cual es contradictorio. Por lo tanto, $(\mathrm{R}_w^\varphi)_{\langle \psi_1, \psi_2 \rangle}(v_1) \subseteq [\![\psi_2]\!]^{\mathfrak{N}_w^\varphi}$ necesariamente. Luego, $[\![\psi_1]\!]^{\mathfrak{N}_w^\varphi} \subseteq \mathrm{SE}^{\mathfrak{N}_w^\varphi}(\langle \psi_1, \psi_2 \rangle)$ y $(\mathrm{R}_w^\varphi)_{\langle \psi_1, \psi_2 \rangle}([\![\psi_1]\!]^{\mathfrak{N}_w^\varphi}) \subseteq [\![\psi_2]\!]^{\mathfrak{N}_w^\varphi}$, y se tiene que $\mathfrak{N}_w^\varphi, v \models \mathsf{Kc}_i(\psi_1, \psi_2)$.

  Si $\mathfrak{N}_w^\varphi, v \models \mathsf{Kc}_i(\psi_1, \psi_2)$, hay dos posibilidades:
  - $[\![\psi_1]\!]^{\mathfrak{N}_w^\varphi} = \emptyset$: Se tiene que $[\![\psi_1]\!]^{\mathfrak{N}^\Gamma} = \emptyset$ necesariamente. Caso contrario, independientemente que $\mathfrak{N}^\Gamma, v \models \mathsf{Kc}_i(\psi_1, \psi_2)$ se cumpla o no, por definición de sel_w^φ, Ítems 5 ó 6, e hipótesis inductiva, se tiene que $[\![\psi_1]\!]^{\mathfrak{N}_w^\varphi} \neq \emptyset$, contradiciendo la hipótesis. Cualquier $\pi \in \mathrm{U}^\Gamma(i)$ funciona como testigo.

- $[\![\psi_1]\!]^{\mathfrak{N}^\varphi_w} \neq \emptyset$: Por hipótesis inductiva, $[\![\psi_1]\!]^{\mathfrak{N}^\Gamma} \neq \emptyset$. Para algún $\pi' \in (\mathrm{U}^\varphi_w)(i)$, $[\![\psi_1]\!]^{\mathfrak{N}^\varphi_w} \subseteq \mathrm{SE}^{\mathfrak{N}^\varphi_w}(\pi')$ y $(\mathrm{R}^\varphi_w)_{\pi'}([\![\psi_1]\!]^{\mathfrak{N}^\varphi_w}) \subseteq [\![\psi_2]\!]^{\mathfrak{N}^\varphi_w}$. Supongamos que $\mathfrak{N}^\Gamma, v \not\models \mathsf{Kc}_i(\psi_1, \psi_2)$. Esto implica que para todo $\pi \in \mathrm{U}^\Gamma(i)$, $[\![\psi_1]\!]^{\mathfrak{N}^\Gamma} \not\subseteq \mathrm{SE}^{\mathfrak{N}^\Gamma}(\pi)$ o $\mathrm{R}^\Gamma_\pi([\![\psi_1]\!]^{\mathfrak{N}^\Gamma}) \not\subseteq [\![\psi_2]\!]^{\mathfrak{N}^\Gamma}$. Más aún, por definición de Act_φ, para todo $\pi = \{a\} \in (\mathrm{U}^\varphi_w)(i)$, con $a \in \mathsf{Act}_\varphi$, $[\![\psi_1]\!]^{\mathfrak{N}^\Gamma} \not\subseteq \mathrm{SE}^{\mathfrak{N}^\Gamma}(a)$ o $\mathrm{R}^\Gamma_a([\![\psi_1]\!]^{\mathfrak{N}^\Gamma}) \not\subseteq [\![\psi_2]\!]^{\mathfrak{N}^\Gamma}$. Por la definición de sel^φ_w, Ítem 6, agregamos los estados para cada $a \in \mathsf{Act}_\varphi$. Sea $\pi' \in (\mathrm{U}^\varphi_w)(i)$. Si $\pi' = \emptyset$, se tiene que $\emptyset \neq [\![\psi_1]\!]^{\mathfrak{N}^\varphi_w} \not\subseteq \mathrm{SE}^{\mathfrak{N}^\varphi_w}(\pi') = \emptyset$, una contradicción. Si $\pi' = \{a\}$ tal que $a \in \mathsf{Act}_\varphi$, sea $w_1' \in [\![\psi_1]\!]^{\mathfrak{N}^\varphi_w} \subseteq [\![\psi_1]\!]^{\mathfrak{N}^\Gamma}$ el estado seleccionado. Si $w_1' \notin \mathrm{SE}^{\mathfrak{N}^\Gamma}(\{a\})$, entonces $\mathrm{R}^\Gamma_a(w_1') = \emptyset$ y $(\mathrm{R}^\varphi_w)_a(w_1') = \emptyset$. Por lo tanto, $w_1' \notin \mathrm{SE}^{\mathfrak{N}^\varphi_w}(\{a\})$. Por otro lado, si existe $w_2 \in \mathrm{R}^\Gamma_a(w_1')$ tal que $w_2 \notin [\![\psi_2]\!]^{\mathfrak{N}^\Gamma}$, por la definición de sel^φ_w e hipótesis inductiva, existe $w_2' \in \mathrm{S}^\varphi_w$ tal que $w_2' \in \mathrm{R}^\Gamma_a(w_1')$ y $w_2' \notin [\![\psi_2]\!]^{\mathfrak{N}^\Gamma}$. Luego, $w_2' \notin [\![\psi_2]\!]^{\mathfrak{N}^\varphi_w}$, y por ende, existe $w_2' \in (\mathrm{R}^\varphi_w)_a(w_1')$ tal que $w_2' \notin [\![\psi_2]\!]^{\mathfrak{N}^\varphi_w}$. En cualquier caso, esto lleva a que $\mathfrak{N}^\varphi_w, v \not\models \mathsf{Kc}_i(\psi_1, \psi_2)$, una contradicción. Con esto, $\mathfrak{N}^\Gamma, v \models \mathsf{Kc}_i(\psi_1, \psi_2)$.

- Caso $\psi = \mathsf{N}(\psi_1, \psi_2)$: Si $\mathfrak{N}^\Gamma, v \models \mathsf{N}(\psi_1, \psi_2)$, hay dos posibilidades:

  - $[\![\psi_1]\!]^{\mathfrak{N}^\Gamma} = \emptyset$: Por hipótesis inductiva, $[\![\psi_1]\!]^{\mathfrak{N}^\varphi_w} = \emptyset$ y basta con elegir $\langle \top, \top \rangle$ como testigo normativo.
  - $[\![\psi_1]\!]^{\mathfrak{N}^\Gamma} \neq \emptyset$: Por el Lema 10.1 y la definición de Act_φ, se tiene que $\langle \psi_1, \psi_2 \rangle$ es un testigo normativo para $\mathsf{N}(\psi_1, \psi_2)$ en $\mathfrak{N}^\Gamma$ y $\langle \psi_1, \psi_2 \rangle \in \mathsf{Act}_\varphi$. Como $[\![\psi_1]\!]^{\mathfrak{N}^\Gamma} \neq \emptyset$, existen $w_1, w_2 \in \mathrm{S}^\Gamma$ tales que $(w_1, w_2) \in \mathrm{R}^\Gamma_{\langle \psi_1, \psi_2 \rangle}$.
    Por definición de $\mathfrak{N}^\varphi_w$, $\langle \psi_1, \psi_2 \rangle \in \mathrm{N}^\varphi_w$ y $(\mathrm{R}^\varphi_w)_{\langle \psi_1, \psi_2 \rangle}$ está definida. Además, por definición de sel^φ_w, Ítem 5, existen $w_1', w_2' \in \mathrm{S}^\varphi_w$ tales que $(w_1', w_2') \in (\mathrm{R}^\varphi_w)_{\langle \psi_1, \psi_2 \rangle}$. Sea $v_1 \in [\![\psi_1]\!]^{\mathfrak{N}^\varphi_w} \subseteq [\![\psi_1]\!]^{\mathfrak{N}^\Gamma}$. Luego, $v_1 \in \mathrm{SE}^{\mathfrak{N}^\Gamma}(\langle \psi_1, \psi_2 \rangle)$ y $\mathrm{R}^\Gamma_{\langle \psi_1, \psi_2 \rangle}(v_1) \subseteq [\![\psi_2]\!]^{\mathfrak{N}^\Gamma}$. Para todo $v_2 \in \mathrm{R}^\Gamma_{\langle \psi_1, \psi_2 \rangle}(v_1)$ se tiene que $v_2 \in [\![\psi_2]\!]^{\mathfrak{N}^\Gamma}$, particularmente $v_2 = w_2'$. Con esto, $w_2' \in (\mathrm{R}^\varphi_w)_{\langle \psi_1, \psi_2 \rangle}(v_1)$ y $v_1 \in \mathrm{SE}^{\mathfrak{N}^\varphi_w}(\{\langle \psi_1, \psi_2 \rangle\})$.
    Supongamos que $(\mathrm{R}^\varphi_w)_{\langle \psi_1, \psi_2 \rangle}(v_1) = \mathrm{R}^\Gamma_{\langle \psi_1, \psi_2 \rangle}(v_1) \cap \mathrm{S}^\varphi_w \not\subseteq [\![\psi_2]\!]^{\mathfrak{N}^\varphi_w}$. Sea $v_2 \in (\mathrm{R}^\varphi_w)_{\langle \psi_1, \psi_2 \rangle}(v_1)$ tal que $v_2 \notin [\![\psi_2]\!]^{\mathfrak{N}^\varphi_w}$, se tiene que $(\mathrm{R}^\varphi_w)_{\langle \psi_1, \psi_2 \rangle}(v_1) \subseteq \mathrm{R}^\Gamma_{\langle \psi_1, \psi_2 \rangle}(v_1)$, y por hipótesis inductiva $v_2 \notin [\![\psi_2]\!]^{\mathfrak{N}^\Gamma}$. Con esto, $\langle \psi_1, \psi_2 \rangle$ no es un testigo válido en $\mathfrak{N}^\Gamma$, lo cual es contradictorio. Por lo tanto, $(\mathrm{R}^\varphi_w)_{\{\langle \psi_1, \psi_2 \rangle\}}(v_1) \subseteq [\![\psi_2]\!]^{\mathfrak{N}^\varphi_w}$ necesariamente. Luego, $[\![\psi_1]\!]^{\mathfrak{N}^\varphi_w} \subseteq \mathrm{SE}^{\mathfrak{N}^\varphi_w}(\langle \psi_1, \psi_2 \rangle)$ y $(\mathrm{R}^\varphi_w)_{\langle \psi_1, \psi_2 \rangle}([\![\psi_1]\!]^{\mathfrak{N}^\varphi_w}) \subseteq [\![\psi_2]\!]^{\mathfrak{N}^\varphi_w}$, y se tiene que $\mathfrak{N}^\varphi_w, v \models \mathsf{N}(\psi_1, \psi_2)$.

  Si $\mathfrak{N}^\varphi_w, v \models \mathsf{N}(\psi_1, \psi_2)$, hay dos posibilidades:

  - $[\![\psi_1]\!]^{\mathfrak{N}^\varphi_w} = \emptyset$: Necesariamente $[\![\psi_1]\!]^{\mathfrak{N}^\Gamma} = \emptyset$. Caso contrario, independientemente que $\mathfrak{N}^\Gamma, v \models \mathsf{N}(\psi_1, \psi_2)$ se cumpla o no, por definición de sel^φ_w, Ítems 5 ó 7, e hipótesis inductiva, se tiene que $[\![\psi_1]\!]^{\mathfrak{N}^\varphi_w} \neq \emptyset$, contradiciendo la hipótesis. Luego, cualquier $\sigma \in \mathrm{N}^\Gamma$ funciona como testigo.
  - $[\![\psi_1]\!]^{\mathfrak{N}^\varphi_w} \neq \emptyset$: Por hipótesis inductiva $[\![\psi_1]\!]^{\mathfrak{N}^\Gamma} \neq \emptyset$. Para algún $\sigma' \in \mathrm{N}^\varphi_w$, $[\![\psi_1]\!]^{\mathfrak{N}^\varphi_w} \subseteq \mathrm{SE}^{\mathfrak{N}^\varphi_w}(\sigma')$ y $(\mathrm{R}^\varphi_w)_{\sigma'}([\![\psi_1]\!]^{\mathfrak{N}^\varphi_w}) \subseteq [\![\psi_2]\!]^{\mathfrak{N}^\varphi_w}$. Supongamos que $\mathfrak{N}^\Gamma, v \not\models \mathsf{N}(\psi_1, \psi_2)$. Esto implica que para todo $\sigma \in \mathrm{N}^\Gamma$, $[\![\psi_1]\!]^{\mathfrak{N}^\Gamma} \not\subseteq \mathrm{SE}^{\mathfrak{N}^\Gamma}(\sigma)$ o $\mathrm{R}^\Gamma_\sigma([\![\psi_1]\!]^{\mathfrak{N}^\Gamma}) \not\subseteq [\![\psi_2]\!]^{\mathfrak{N}^\Gamma}$. Más aún, por definición de Act_φ se tiene que para todo $\sigma' = \langle \psi_1, \psi_2 \rangle \in \mathrm{N}^\varphi_w$, $[\![\psi_1]\!]^{\mathfrak{N}^\Gamma} \not\subseteq \mathrm{SE}^{\mathfrak{N}^\Gamma}(\sigma')$ o $\mathrm{R}^\Gamma_{\sigma'}([\![\psi_1]\!]^{\mathfrak{N}^\Gamma}) \not\subseteq [\![\psi_2]\!]^{\mathfrak{N}^\Gamma}$. Por la definición de sel^φ_w, Ítem 7, agregamos los estados para cada $\sigma' \in \mathsf{Act}_\varphi$. Sean $\sigma' \in \mathrm{N}^\varphi_w$ y $w_1' \in [\![\psi_1]\!]^{\mathfrak{N}^\varphi_w} \subseteq [\![\psi_1]\!]^{\mathfrak{N}^\Gamma}$ el estado seleccionado. Si $w_1' \notin \mathrm{SE}^{\mathfrak{N}^\Gamma}(\sigma')$, $\mathrm{R}^\Gamma_{\sigma'}(w_1') = \emptyset$, entonces $(\mathrm{R}^\varphi_w)_{\sigma'}(w_1') = \emptyset$ y por lo tanto $w_1' \notin \mathrm{SE}^{\mathfrak{N}^\varphi_w}(\sigma')$. Por otro lado, si existe $w_2 \in \mathrm{R}^\Gamma_a(w_1')$ tal que $w_2 \notin [\![\psi_2]\!]^{\mathfrak{N}^\Gamma}$, por la definición de sel^φ_w e hipótesis inductiva, existe $w_2' \in \mathrm{S}^\varphi_w$ tal que $w_2' \in \mathrm{R}^\Gamma_a(w_1')$ y $w_2' \notin [\![\psi_2]\!]^{\mathfrak{N}^\varphi_w}$ Luego, existe $w_2' \in (\mathrm{R}^\varphi_w)_a(w_1')$ tal que $w_2' \notin [\![\psi_2]\!]^{\mathfrak{N}^\varphi_w}$. En cualquier caso, esto lleva a que $\mathfrak{N}^\varphi_w, v \not\models \mathsf{N}(\psi_1, \psi_2)$, una contradicción. Con esto, $\mathfrak{N}^\Gamma, v \models \mathsf{N}(\psi_1, \psi_2)$.

Queda demostrar que $\mathfrak{N}^\varphi_w$ es de tamaño polinomial con respecto al tamaño de la fórmula φ. La función de selección agrega estados de $\mathfrak{N}^\Gamma$ sólo por cada subfórmula de φ del tipo Kc_i o N. El número de estados agregados en cada caso es polinomial en el tamaño de φ. Por lo tanto, el

Algoritmo 2: Model Checking para Kc_i

Datos: $\mathfrak{N}$ un NLTS^U, $w \in \mathrm{D}_{\mathfrak{N}}$, $\mathsf{Kc}_i(\psi, \varphi)$
Resultado: kc
$lab(\mathsf{Kc}_i(\psi, \varphi)) \leftarrow \emptyset$;
para todo $\pi \in \mathrm{U}(i), \pi \subseteq \mathrm{N}$ **hacer**
  $kc \leftarrow Verdadero$;
  si $\pi == \emptyset$ **entonces**
    $kc \leftarrow lab(\psi) == \emptyset$;
  fin
  para todo $\sigma \in \pi$ **hacer**
    para todo $v \in lab(\psi)$ **hacer**
      $kc \leftarrow (kc \ \& \ v \in \mathrm{SE}(\sigma) \ \& \ \mathrm{R}_\sigma(v) \subseteq lab(\varphi))$;
    fin
  fin
  si kc **entonces**
    $lab(\mathsf{Kc}_i(\psi, \varphi)) \leftarrow \mathrm{S}$;
  fin
fin

Algoritmo 3: Model Checking para N

Datos: $\mathfrak{N}$ un NLTS^U, $w \in \mathrm{D}_{\mathfrak{N}}$, $\mathsf{N}(\psi, \varphi)$
Resultado: nd
$lab(\mathsf{N}(\psi, \varphi)) \leftarrow \emptyset$;
para todo $\pi \in \mathrm{N}$ **hacer**
  $nd \leftarrow Verdadero$;
  para todo $\sigma \in \pi$ **hacer**
    para todo $v \in lab(\psi)$ **hacer**
      $nd \leftarrow (nd \ \& \ v \in \mathrm{SE}(\sigma) \ \& \ \mathrm{R}_\sigma(v) \subseteq lab(\varphi))$;
    fin
  fin
  si nd **entonces**
    $lab(\mathsf{N}(\psi, \varphi)) \leftarrow \mathrm{S}$;
  fin
fin

tamaño de S_w^φ es polinomial. Como $(\mathrm{U}_w^\varphi)(i)$ y N_w^φ también son polinomiales, el tamaño de $\mathfrak{N}_w^\varphi$ es polinomial al tamaño de φ. $\square$

Para demostrar que el problema de satisfacibilidad está en NP, basta con mostrar que el problema de model checking está en P.

Proposición 10.12. *El problema de model checking para* $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N}}$ *está en* P.

Demostración. Sean un NLTS^U $\mathfrak{N}$, un estado $w \in \mathrm{D}_{\mathfrak{N}}$ y una fórmula φ, se define un algoritmo de etiquetado de abajo hacia arriba (bottom-up) en tiempo polinomial que verifica si $\mathfrak{N}, w \models \varphi$, siguiendo las ideas de la lógica modal básica ML [26]. Con esto, se tienen los Algoritmos 2 y 3, que computan el valor de verdad para las fórmulas Kc_i y N respectivamente, sobre un NLTS^U $\mathfrak{N} = \langle \mathrm{S}, \mathrm{R}, \mathrm{U}, \mathrm{V} \rangle$ y $w \in \mathrm{S}$.

Como $\mathrm{U}(i)$ es no vacío, el primer ciclo **para todo** se ejecuta necesariamente. Si $\pi = \emptyset$, basta con verificar si $lab(\psi) = \emptyset$. Si $\pi \neq \emptyset$ pero $lab(\psi) = \emptyset$, entonces la fórmula $\mathsf{Kc}_i(\psi, \varphi)$ es trivialmente verdadera. En caso contrario, kc se mantendrá verdadera sólo si las condiciones apropiadas para la satisfacibilidad de $\mathsf{Kc}_i(\psi, \varphi)$ se cumplen para algún $\pi \in \mathrm{U}(i)$. Si no, la inicialización de $lab(\mathsf{Kc}_i(\psi, \varphi))$ como $\emptyset$ no se sobrescribe. Las condiciones $\pi \subseteq \mathrm{N}$, $lab(\psi) = \emptyset$, $v \in \mathrm{SE}(\sigma)$ y R_σ se pueden verificar en tiempo polinomial. El algoritmo para el operador N es similar y se ejecuta en tiempo polinomial. Por lo tanto, el problema de model checking está en P. $\square$

Con esto, caracterizamos la complejidad del problema de satisfacibilidad para $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N}}$.

Teorema 10.2. *El problema de satisfacibilidad para* $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N}}$ *es* NP*-completo.*

Demostración. Hardness para NP (es decir, la cota inferior) se sigue de que la lógica proposicional, un fragmento de $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N}}$, es NP-completo [12, 35]. Por la Proposición 10.11, cada fórmula satisfacible φ tiene un modelo de tamaño polinomial con respecto a φ. Con esto, se puede adivinar un modelo polinomial $\mathfrak{N}$ y un estado $w \in \mathrm{D}_{\mathfrak{N}}$, y verificar $\mathfrak{N}, w \models \varphi$ en tiempo polinomial, por la Proposición 10.12. Por lo tanto, el problema de satisfacibilidad está en NP. □

10.5. Razonando sobre habilidades

En las secciones anteriores de este capítulo se ha trabajado con un formalismo que permite expresar y razonar sobre cursos de acción normativos y las responsabilidades de los agentes. Estas nociones están representadas sintácticamente en $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N}}$ por las modalidades N y Kc_i respectivamente. Desde un punto de vista semántico, las mismas modalidades son interpretadas en un modelo relativo a un conjunto N de planes normativos. Sin embargo, esta lógica no permite hablar directamente sobre las habilidades de un agentes. Es decir, lo que el agente puede hacer en términos de planes, independientemente de las normas establecidas y su propia percepción.

Contar con una modalidad que describa esta característica termina siendo útil a la hora de estudiar las interacciones entre las habilidades y lo que los agentes hacen de acuerdo con las normas. Para ello, en esta sección se analiza el impacto de agregar una nueva modalidad para este aspecto, denotada como S, a $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N}}$. La lógica resultante se refiere como $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N},\mathsf{S}}$.

Definición 10.9. Sean Prop y Act conjuntos numerables de símbolos proposicionales y acciones básicas respectivamente, y Agt un conjunto no vacío de nombres de agentes. Las fórmulas de $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N},\mathsf{S}}$ son dadas por la siguiente gramática

$$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid \mathsf{S}(\psi,\varphi) \mid \mathsf{N}(\varphi,\varphi) \mid \mathsf{Kc}_i(\varphi,\varphi),$$

con $p \in \mathsf{Prop}$ y $i \in \mathsf{Agt}$. Las constantes booleanas y demás conectores $\bot$, $\top$, $\varphi \wedge \psi$, $\varphi \to \psi$, y $\varphi \leftrightarrow \psi$ se definen de la forma usual. Intuitivamente, la fórmula $\mathsf{S}(\psi,\varphi)$ se lee como *"existe un curso de acción que cumple φ dado ψ"*.

Notar que la semántica de S es exactamente la semántica vista en el Capítulo 3 para la modalidad Kh [116,117]. Si bien en el Capítulo 4 y en [8] se argumentaba que la definición de este operador asumía un cierto nivel de omnisciencia por parte del agente, esto no genera problemas para $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N},\mathsf{S}}$, ya que S no es una modalidad *epistémica*, sino una modalidad de *habilidad*.

Definición 10.10 (S en NLTS^Us)**.** Sea $\mathfrak{N} = \langle \mathrm{S}, \mathrm{R}, \mathrm{V}, \mathrm{U}, \mathrm{N}\rangle$ un NLTS^U sobre Act, Prop y Agt, $w \in \mathrm{S}$. Se extiende la definición de la Definición 10.3 para el operador S como:

$$\mathfrak{N}, w \models \mathsf{S}(\psi,\varphi) \quad \text{sii} \quad \text{existe } \sigma \in \mathsf{Act}^* \text{ tal que } \textbf{(S-1)}\ [\![\psi]\!]^{\mathfrak{N}} \subseteq \mathrm{SE}(\sigma), \text{ y } \textbf{(S-2)}\ \mathrm{R}_\sigma([\![\psi]\!]^{\mathfrak{N}}) \subseteq [\![\varphi]\!]^{\mathfrak{N}}$$

El plan σ que hace verdadero $\mathsf{S}(\psi,\varphi)$ en $\mathfrak{N}$ es llamado *testigo* para $\mathsf{S}(\psi,\varphi)$ en $\mathfrak{N}$.

Ejemplo 10.4. Teniendo en cuenta el procedimiento de emergencia, la fórmula $\mathsf{S}(f \wedge c, s)$ establece que "existe un curso de acción el cual siempre llega a un lugar seguro (s) en caso de incendio (f), teniendo en cuenta que se es capaz de hacerlo (c)", mientras que $\mathsf{S}(\top, s)$ expresa que "existe un curso de acción que siempre ($\top$) permite a los agentes alcanzar un lugar seguro (s)". En el contexto del LTS dado en el Ejemplo 10.1, se tiene que

- $\mathfrak{N}, w_1 \models \mathsf{S}(f \wedge c, s)$ y
- $\mathfrak{N}, w_1 \not\models \mathsf{S}(\top, s)$.

Para la primera fórmula, el plan $\sigma_6 = \mathsf{use.ramp}$ se puede utilizar como testigo, aún si este no cumple con las normativas de activar la alarma y llamar a los bomberos. Notar que la segunda fórmula no se cumple dado que en el estado w_2 no existe un plan que nos conduzca siempre a un lugar seguro (s). Intuitivamente, el LTS de este ejemplo puede verse como una "ilustración" de las acciones ejecutables en caso de un incendio. El primer ítem habla que en ese escenario es posible llegar a un lugar seguro si se es capaz de hacerlo. Mientras tanto, el segundo ítem establece que no siempre es posible alcanzar un lugar seguro si, por ejemplo, estamos atrapados en un ascensor. □

Bloque $\mathcal{L}^+$:	$\mathsf{4SA}$	$\vdash \mathsf{S}(\psi,\varphi) \to \mathsf{AS}(\psi,\varphi)$
	$\mathsf{5SA}$	$\vdash \neg\mathsf{S}(\psi,\varphi) \to \mathsf{A}\neg\mathsf{S}(\psi,\varphi)$
	EMPS	$\vdash \mathsf{A}(\psi \to \varphi) \to \mathsf{S}(\psi,\varphi)$
	COMPS	$\vdash (\mathsf{S}(\psi,\chi) \wedge \mathsf{S}(\chi,\varphi)) \to \mathsf{S}(\psi,\varphi)$
Bloque $\mathcal{L}^+_{\mathrm{NLTS}^U}$:	NS	$\vdash \mathsf{N}(\psi,\varphi) \to \mathsf{S}(\psi,\varphi)$
	$\mathsf{SN}\bot$	$\vdash \mathsf{S}(\psi,\bot) \to \mathsf{N}(\psi,\bot)$

Tabla 10.2: Axiomas adicionales $\mathcal{L}_{\mathsf{S}}$ para $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N},\mathsf{S}}$.

Para finalizar con esta sección, presentaremos una axiomatización correcta y fuertemente completa para $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N},\mathsf{S}}$. La misma es una extensión de la vista para $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N}}$ (Tabla 10.1) que cuenta con los siguientes axiomas (Tabla 10.2):

1. $\mathsf{4SA}$ y $\mathsf{5SA}$: Describen que S es una modalidad global de la misma forma que lo hacen $\mathsf{4KcA}$, $\mathsf{5KcA}$, $\mathsf{4NA}$ y $\mathsf{5NA}$ para Kc_i y N.
2. EMPS y COMPS: Son los axiomas EMP y COMP de la Tabla 3.1. El primero habla que es posible convertir implicaciones universalmente válidas en habilidades sin realizar ninguna acción (atestiguada por el plan vacío ϵ). Mientras tanto, el segundo especifica que los cursos de acción que tienen un objetivo o un contexto en común pueden componerse.
3. NS: Establece la interacción entre N y S. Caracteriza que todo lo que esté regulado por normas es también factible. Esto descarta sistemas normativos que cuentan con normas imposibles. Por ejemplo, si pensamos las normas en términos de obligaciones, este axioma nos dice que la lógica se adhiere al principio *impossibilium nulla obligatio est*, que establece que las normas imposibles son rechazadas. Como se puede ver, la recíproca de NS no es un teorema en la lógica ya que es posible que ciertos cursos de acción no estén regulados por las normas.
4. $\mathsf{SN}\bot$: Análogamente a $\mathsf{NKc}\bot$, vincula las globalidades de A, definida en términos de N, con Kc_i. Dado que las tres modalidades son globales, resulta indistinto definir A en términos de Kc_i, N, S o una combinación de estas tres. Por lo tanto, A sigue definida en términos de N.

Extendiendo los conceptos de deducción, consistencia y maximalidad de la Definición 10.4 para el sistema axiomático $\mathcal{L}_{\mathsf{Kc}_i,\mathsf{N},\mathsf{S}} = \mathcal{L}_{\mathsf{Kc}_i,\mathsf{N}} + \mathcal{L}_{\mathsf{S}}$, se derivan los siguientes teoremas.

Proposición 10.13. *Las siguientes fórmulas son derivables usando los axiomas y reglas en las Tablas 10.1 y 10.2:*

- $\vdash (\mathsf{A}(\psi \to \chi) \wedge \mathsf{S}(\chi,\rho) \wedge \mathsf{A}(\rho \to \varphi)) \to \mathsf{S}(\psi,\varphi)$ *(SA)*
- $\vdash (\mathsf{E}\psi \wedge \mathsf{S}(\psi,\varphi)) \to \mathsf{E}\varphi$ *(SE)*

Demostración. La demostración de SA es análoga a la vista en la Proposición 6.2 para Kh. Para el caso de SE, se tiene la Proposición 5.1 para Kh_i usando SA. □

Además, se definen los elementos del modelo canónico de $\mathcal{L}_{\mathsf{Kc}_i,\mathsf{N},\mathsf{S}}$ considerando la nueva modalidad S.

Definición 10.11. Sea $\mathbf{\Phi}$ el conjunto de todos los conjuntos maximales consistentes con $\mathcal{L}_{\mathsf{Kc}_i,\mathsf{N},\mathsf{S}}$. Para todo $\Gamma \in \mathbf{\Phi}$, se extiende la definición de Definición 10.11 define:

$$\begin{array}{llll} \Gamma|_{\mathsf{N}} & := \{\mathsf{N}(\psi,\varphi) \mid \mathsf{N}(\psi,\varphi) \in \Gamma\}, & \Gamma|_{\mathsf{A}} & := \{\mathsf{A}\varphi \mid \mathsf{A}\varphi \in \Gamma\}, \\ \Gamma|_{\mathsf{Kc}_i} & := \{\mathsf{Kc}_i(\psi,\varphi) \mid \mathsf{Kc}_i(\psi,\varphi) \in \Gamma\}, & \Gamma|_{\mathsf{Kc}} & := \bigcup_{i \in \mathsf{Agt}} \Gamma|_{\mathsf{Kc}_i}, \\ \Gamma|_{\mathsf{S}} & := \{\mathsf{S}(\psi,\varphi) \mid \mathsf{S}(\psi,\varphi) \in \Gamma\}, & \mathsf{Act}^{\Gamma} & := \{\langle \psi,\varphi \rangle \mid \mathsf{S}(\psi,\varphi) \in \Gamma\}. \end{array}$$

Definición 10.12. Sea $\Gamma \in \mathbf{\Phi}$ un conjunto maximal consistente fórmulas, el *modelo canónico* para Γ es una tupla $\mathfrak{N}^{\Gamma} = \langle \mathrm{S}^{\Gamma}, \mathrm{R}^{\Gamma}, \mathrm{V}^{\Gamma}, \mathrm{U}^{\Gamma}, \mathrm{N}^{\Gamma} \rangle$ sobre Act^{Γ}, Agt y Prop donde:

- $\mathrm{S}^{\Gamma} = \{\Delta \in \mathbf{\Phi} \mid \Delta|_{\mathsf{A}} = \Gamma|_{\mathsf{A}}\}$;

- $\mathrm{R}^{\Gamma}_{\langle\psi,\varphi\rangle} = \{(\Delta_1,\Delta_2) \in \mathrm{S}^{\Gamma} \times \mathrm{S}^{\Gamma} \mid \mathsf{S}(\psi,\varphi) \in \Gamma, \psi \in \Delta_1, \varphi \in \Delta_2\}$;
- $\mathrm{R}^{\Gamma} = \{\mathrm{R}^{\Gamma}_{\langle\psi,\varphi\rangle} \mid \mathsf{S}(\psi,\varphi) \in \Gamma\}$;
- $\mathrm{V}^{\Gamma}(\Delta) = \{p \in \mathsf{Prop} \mid p \in \Delta\}$;
- $\mathrm{U}^{\Gamma}(i) = \{\{\langle\psi,\varphi\rangle\} \mid \mathsf{Kc}_i(\psi,\varphi) \in \Gamma\} \cup \{\emptyset\}$; y
- $\mathrm{N}^{\Gamma} = \{\langle\psi,\varphi\rangle \mid \mathsf{N}(\psi,\varphi) \in \Gamma\}$.

Con esto, procedemos a mostrar las propiedades de dicho modelo. En la mayoría de estas, la demostración es similar a las vistas en la Sección 10.3.

Proposición 10.14. *$\mathfrak{N}^{\Gamma} = \langle \mathrm{S}^{\Gamma}, \mathrm{R}^{\Gamma}, \mathrm{V}^{\Gamma}, \mathrm{U}^{\Gamma}, \mathrm{N}^{\Gamma}\rangle$ es un NLTS^U. Además, $\{\langle\psi,\varphi\rangle\} \in \mathrm{U}^{\Gamma}(i)$ implica $\langle\psi,\varphi\rangle \in \mathrm{N}^{\Gamma}$, y $\mathrm{R}^{\Gamma}_{\langle\psi,\varphi\rangle} \neq \emptyset$ implica $\mathsf{S}(\psi,\varphi) \in \Gamma$.*

Demostración. Análoga a las Proposiciones 10.4 y 10.3. □

Proposición 10.15. *Sean $\Delta_1, \Delta_2 \in \mathrm{S}^{\Gamma}$ y $\mathsf{X} \in \{\mathsf{Kc}_i, \mathsf{N}, \mathsf{S}\}$; se tiene que $\Delta_1|_{\mathsf{X}} = \Delta_2|_{\mathsf{X}} = \Gamma|_{\mathsf{X}}$.*

Demostración. Análoga a la Proposición 10.5 usando además 4SA y 5SA. □

Proposición 10.16. *Si $\mathrm{R}^{\Gamma}_{\langle\psi,\varphi\rangle}(\Delta) \neq \emptyset$, entonces, para todo $\Delta' \in \mathrm{S}^{\Gamma}$, $\varphi \in \Delta'$ implica $\Delta' \in \mathrm{R}^{\Gamma}_{\langle\psi,\varphi\rangle}(\Delta)$.*

Demostración. Análoga a la Proposición 10.6 usando que $\mathsf{S}(\psi,\varphi) \in \Gamma$. □

Proposición 10.17. *Sea φ una fórmula de $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N}}$. Si $\varphi \in \Delta$ para todo $\Delta \in \mathrm{S}^{\Gamma}$, entonces $\mathsf{A}\varphi \in \Delta$ para todo $\Delta \in \mathrm{S}^{\Gamma}$.*

Demostración. La prueba es análoga a la Proposición 9.5, extendiendo la Definición 10.11 con $\Delta|_{\neg\mathsf{A}} = \{\neg\mathsf{A}\varphi \mid \neg\mathsf{A}\varphi \in \Delta\}$ y usando 4NA y 5NA. □

Proposición 10.18. *Sean ψ, ψ', φ' fórmulas de $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N}}$. Si para todo $\Delta \in \mathrm{S}^{\Gamma}$ tal que $\psi \in \Delta$, tiene un sucesor vía $\mathrm{R}^{\Gamma}_{\langle\psi',\varphi'\rangle}$. Entonces $\mathsf{A}(\psi \to \psi') \in \Delta$ para todo $\Delta \in \mathrm{S}^{\Gamma}$.*

Demostración. Análoga a la Proposición 5.6 usando la Proposición 10.17. □

Proposición 10.19. *Sea $\mathsf{X} \in \{\mathsf{Kc}_i, \mathsf{N}, \mathsf{S}\}$, si existe $\Delta \in \mathrm{S}^{\Gamma}$ tal que $\{\psi, \mathsf{X}(\psi,\varphi)\} \subseteq \Delta$, entonces existe $\Delta' \in \mathrm{S}^{\Gamma}$ tal que $\varphi \in \Delta'$.*

Demostración. Análoga a la Proposición 10.9 usando las Proposiciones 10.17, 10.13 y 10.2. □

Con estas propiedades, procedemos al "truth lemma" para $\mathfrak{N}^{\Gamma}$.

Lema 10.2. *Sean $\Gamma \in \mathbf{\Phi}$ y $\mathfrak{N}^{\Gamma} = \langle \mathrm{S}^{\Gamma}, \mathrm{R}^{\Gamma}, \mathrm{V}^{\Gamma}, \mathrm{U}^{\Gamma}, \mathrm{N}^{\Gamma}\rangle$. Para todo $\Theta \in \mathrm{S}^{\Gamma}$ y φ fórmula de $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N},\mathsf{S}}$,*

$$\mathfrak{N}^{\Gamma}, \Theta \models \varphi \qquad \textit{si y sólo si} \qquad \varphi \in \Theta.$$

Demostración. Sea $\mathfrak{N}^{\Gamma} = \langle \mathrm{S}^{\Gamma}, \mathrm{R}^{\Gamma}, \mathrm{V}^{\Gamma}, \mathrm{U}^{\Gamma}, \mathrm{N}^{\Gamma}\rangle$. La prueba es por inducción en φ. Los casos atómicos y para los operadores booleanos son directos. Por otro lado, los casos Kc_i y N son análogos al Lema 10.1, por lo que queda verificar el caso para S.

Caso $\mathsf{S}(\psi,\varphi)$: ($\Rightarrow$) Supongamos que $\mathfrak{N}^{\Gamma}, \Theta \models \mathsf{S}(\psi,\varphi)$. Luego, por $\models$, existe $\sigma \in (\mathsf{Act}^{\Gamma})^*$ tal que $[\![\psi]\!]^{\mathfrak{N}^{\Gamma}} \subseteq \mathrm{SE}(\sigma)$ y $\mathrm{R}_{\sigma}([\![\psi]\!]^{\mathfrak{N}^{\Gamma}}) \subseteq [\![\varphi]\!]^{\mathfrak{N}^{\Gamma}}$. Consideremos tres casos:

- $[\![\psi]\!]^{\mathfrak{N}^{\Gamma}} = \emptyset$: Por hipótesis inductiva y maximalidad, $\neg\psi \in \Delta$ para todo $\Delta \in \mathrm{S}^{\Gamma}$. Por la Proposición 10.17, $\mathsf{A}\neg\psi \in \Delta$ para todo $\Delta \in \mathrm{S}^{\Gamma}$ y por NS, $\mathsf{S}(\psi,\bot) \in \Delta$. Como $\bot \to \varphi$ y $\psi \to \psi$ son tautologías, por NECA $\mathsf{A}(\bot \to \varphi)$ y $\mathsf{A}(\psi \to \psi)$ están en Δ. Utilizando una instancia de SA en la Proposición 10.13, $(\mathsf{A}(\psi \to \psi) \wedge \mathsf{S}(\psi,\bot) \wedge \mathsf{A}(\bot \to \varphi)) \to \mathsf{S}(\psi,\varphi) \in \Delta$. Por MP, $\mathsf{S}(\psi,\varphi) \in \Delta$ para todo $\Delta \in \mathrm{S}^{\Gamma}$. Luego, $\mathsf{S}(\psi,\varphi) \in \Theta$.
- $[\![\psi]\!]^{\mathfrak{N}^{\Gamma}} \neq \emptyset$ y $\sigma = \epsilon$: Para todo $\Delta \in \mathrm{S}^{\Gamma}$, si $\Delta \in [\![\psi]\!]^{\mathfrak{N}^{\Gamma}}$ entonces $\Delta \in [\![\varphi]\!]^{\mathfrak{N}^{\Gamma}}$. Es decir, para todo $\Delta \in \mathrm{S}^{\Gamma}$, $\Delta \notin [\![\psi]\!]^{\mathfrak{N}^{\Gamma}}$ o $\Delta \in [\![\varphi]\!]^{\mathfrak{N}^{\Gamma}}$. Por hipótesis inductiva, $\psi \notin \Delta$ o $\varphi \in \Delta$. Como Δ es un conjunto maximal consistente, $\psi \to \varphi \in \Delta$ para todo $\Delta \in \mathrm{S}^{\Gamma}$. Por la Proposición 10.17, $\mathsf{A}(\psi \to \varphi) \in \Delta$ para todo $\Delta \in \mathrm{S}^{\Gamma}$ y por EMPS, $\mathsf{S}(\psi,\varphi) \in \Delta$. Luego, $\mathsf{S}(\psi,\varphi) \in \Theta$.

- $[\![\psi]\!]^{\mathfrak{N}^\Gamma} \neq \emptyset$ y $\sigma = \langle\psi_1,\varphi_1\rangle\cdots\langle\psi_n,\varphi_n\rangle$: Primero se probará por inducción la siguiente propiedad:

  $P(k)$: para todo $k \leq n$, ***(1)*** $\mathsf{S}(\psi,\varphi_k) \in \Gamma$ y ***(2)*** cada estado φ_k es alcanzable vía $\langle\psi_1,\varphi_1\rangle\cdots\langle\psi_k,\varphi_k\rangle$ desde algún estado ψ.

  - $P(1)$: Por la semántica de S, cada estado ψ tiene un sucesor vía $\langle\psi_1,\varphi_1\rangle$. Por la Proposición 10.16, todo estado φ_1 puede ser alcanzado desde algún estado ψ vía $\langle\psi_1,\varphi_1\rangle$, demostrando lo segundo. Además, por la Proposición 10.18, se tiene que $\mathsf{A}(\psi \to \psi_1) \in \Delta$, para todo $\Delta \in \mathrm{S}^\Gamma$. Luego, por EMPS, $\mathsf{S}(\psi,\psi_1) \in \Delta$. Como establecimos que $\mathrm{R}^\Gamma_{\langle\psi_1,\varphi_1\rangle} \neq \emptyset$, usando la Proposición 10.14, se cumple que $\mathsf{S}(\psi_1,\varphi_1) \in \Gamma$ y con esto $\mathsf{S}(\psi_1,\varphi_1) \in \Delta$. Por COMPS y la Proposición 10.15, $\mathsf{S}(\psi,\varphi_1) \in \Delta|_{\mathsf{S}} = \Gamma|_{\mathsf{S}} \subseteq \Gamma$, demostrando lo primero.
  - $P(i-1) \to P(i)$: Por hipótesis inductiva, $\mathsf{S}(\psi,\varphi_{i-1}) \in \Gamma$ y todo estado φ_{i-1} es alcanzado desde un estado ψ vía $\langle\psi_1,\varphi_1\rangle\cdots\langle\psi_{i-1},\varphi_{i-1}\rangle$. Como σ es fuertemente ejecutable en todos los estados ψ, cada estado φ_{i-1} tiene un sucesor vía $\mathrm{R}^\Gamma_{\langle\psi_i,\varphi_i\rangle}$. Por la Proposición 10.18, $\mathsf{A}(\varphi_{i-1} \to \psi_i) \in \Gamma$ y EMPS, $\mathsf{S}(\varphi_{i-1},\psi_i) \in \Gamma$. Por definición de R^Γ, $\mathsf{S}(\psi_i,\varphi_i) \in \Gamma$. Por lo tanto, se tiene que $\mathsf{S}(\psi,\varphi_{i-1})$, $\mathsf{S}(\varphi_{i-1},\psi_i)$, $\mathsf{S}(\psi_i,\varphi_i) \in \Gamma$. Usando COMPS dos veces y MP, $\mathsf{S}(\psi,\varphi_i) \in \Gamma$, demostrando lo primero. Como cada estado φ_{i-1} tiene un sucesor vía $\mathrm{R}^\Gamma_{\langle\psi_i,\varphi_i\rangle}$, por la Proposición 10.16, todos los estados φ_i pueden ser alcanzados vía $\mathrm{R}^\Gamma_{\langle\psi_i,\varphi_i\rangle}$ desde estados φ_{i-1}. Luego, por hipótesis inductiva, todos los estados φ_i pueden ser alcanzados vía $\langle\psi_1,\varphi_1\rangle\cdots\langle\psi_i,\varphi_i\rangle$ desde algún estado ψ, lo que completa lo segundo.

  Continuando con la prueba, para $k = n$, la propiedad $P(n)$ indica que ***(1)*** $\mathsf{S}(\psi,\varphi_n) \in \Gamma$ y ***(2)*** cada estado φ_n es alcanzado vía σ desde algún estado ψ. Dado que $\sigma = \langle\psi_1,\varphi_1\rangle\cdots\langle\psi_n,\varphi_n\rangle$ es un testigo de $\mathsf{S}(\psi,\varphi)$, para cada estado $\Delta \in \mathrm{S}^\Gamma$ tal que $\varphi_n \in \Delta$, $\mathfrak{N}^\Gamma, \Delta \models \varphi$. Luego, por hipótesis inductiva, si $\varphi_n \in \Delta$ entonces $\varphi \in \Delta$. Por lo tanto, para cada $\Delta \in \mathrm{S}^\Gamma$, se tiene que $\varphi_n \to \varphi \in \Delta$. Usando la Proposición 10.17, $\mathsf{A}(\varphi_n \to \varphi) \in \Delta$ y EMPS, $\mathsf{S}(\varphi_n,\varphi) \in \Delta|_{\mathsf{S}} \subseteq \Gamma$. Con esto, por COMPS, $\mathsf{S}(\psi,\varphi) \in \Gamma$ y $\mathsf{S}(\psi,\varphi) \in \Theta$.

($\Leftarrow$) Supongamos que $\mathsf{S}(\psi,\varphi) \in \Theta$. Luego, por la Proposición 10.15, $\mathsf{S}(\psi,\varphi) \in \Delta$ para todo $\Delta \in \mathrm{S}^\Gamma$. Más aún, $\mathsf{S}(\psi,\varphi) \in \Gamma$ y $\mathrm{R}^\Gamma_{\langle\psi,\varphi\rangle}$ está definida. Para demostrar que $\mathfrak{N}^\Gamma, \Theta \models \mathsf{S}(\psi,\varphi)$ consideramos dos posibilidades:

- No existe Δ tal que $\psi \in \Delta$: Por hipótesis inductiva, $[\![\psi]\!]^{\mathfrak{N}^\Gamma} = \emptyset$. Usando $\sigma = \epsilon$, se tiene trivialmente que $\mathfrak{N}^\Gamma, \Theta \models \mathsf{S}(\psi,\varphi)$.
- Existe Δ tal que $\psi \in \Delta$: Por la Proposición 10.19, existe Δ' tal que $\varphi \in \Delta'$. Por hipótesis inductiva, $\mathfrak{N}^\Gamma, \Delta \models \psi$ y $\mathfrak{N}^\Gamma, \Delta' \models \varphi$. Como está definido, $\sigma = \langle\psi,\varphi\rangle$ es fuertemente ejecutable en todos los estados ψ (dado que existe un sucesor Δ' vía $\mathrm{R}^\Gamma_{\langle\psi,\varphi\rangle}$) y alcanza desde estos sólo estados φ vía σ (por construcción de $\mathrm{R}^\Gamma_{\langle\psi,\varphi\rangle}$). Con esto, $\mathfrak{N}^\Gamma, \Theta \models \mathsf{S}(\psi,\varphi)$.

$\square$

Con el Lema 10.2, finalmente se obtienen las propiedades de esta axiomatización.

Teorema 10.3. *El sistema axiomático $\mathcal{L}_{\mathsf{Kc}_i,\mathsf{N},\mathsf{S}} = \mathcal{L}_{\mathsf{Kc}_i,\mathsf{N}} + \mathcal{L}_{\mathsf{S}}$ (Tablas 10.1 y 10.2) es correcto y fuertemente completo para $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N},\mathsf{S}}$ con respecto a la clase de todos los NLTS^Us.*

Demostración. Para la correctitud, basta con mostrar que los axiomas del sistema son validos y que sus reglas preservan la validez, que resulta ser directo. Para la completitud fuerte, sea Γ' un conjunto de fórmulas consistente con $\mathcal{L}_{\mathsf{Kc}_i,\mathsf{N},\mathsf{S}}$. Teniendo en cuenta la Observación 5.2 aplicada a $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N},\mathsf{S}}$, Γ' puede ser extendido a un conjunto maximal consistente con $\mathcal{L}_{\mathsf{Kc}_i,\mathsf{N},\mathsf{S}}$ $\Gamma \supseteq \Gamma'$. Por el Lema 10.2, $\mathfrak{N}^\Gamma, \Gamma \models \Gamma'$ y Γ' es satisfacible. Dado que $\mathfrak{N}^\Gamma$ es un NLTS^U (Proposición 10.14), esto termina de demostrar la completitud fuerte. $\square$

Concluimos este capítulo con un comentario sobre el comportamiento computacional de $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N},\mathsf{S}}$. Como se mencionó, la semántica de S es la misma que Kh del Capítulo 3 y de [116,117]. En [72], se muestra que el problema de satisfacibilidad es decidible para esta lógica (es más, el resultado es probado para lógicas más generales). A pesar de que la clase de complejidad del problema de satisfacibilidad para $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N},\mathsf{S}}$ es por el momento un problema abierto, al menos se tiene que los fragmentos $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N}}$ y L_{S} (fórmulas sin las modalidades N y Kc_i) son NP-completo (Teorema 10.2) y a lo sumo $\mathsf{NP}^{\mathsf{NP}}$ (Teorema 3.5) respectivamente.

Parte IV

Conclusiones

Capítulo 11

Revisión

> I just hate sitting and writing — I had to do that in school. Plus, I have terrible handwriting.
>
> Saul Kripke

A lo largo de este trabajo hemos definido una lógica epistémica de saber cómo basada en incertidumbre. En el mismo comenzamos introduciendo los conceptos y definiciones necesarias para el resto de la tesis, como así también el contexto y motivación que nos guiaron en el proceso. En este capítulo haremos un resumen del contenido de la tesis.

En el Capítulo 1, se habló del concepto y la importancia de la representación de la realidad mediante modelos en ciencias de la computación. En particular, nos enfocamos en las lógicas modales, extensiones de la lógica proposicional que cuentan con una flexibilidad en la definición de operadores que representen "modos de verdad". Para dar una noción general de estas lógicas, partimos de la lógica modal básica [24, 26]. Además de la definición de la sintaxis y semántica de la lógica, introdujimos diversas herramientas de utilidad para el estudio de los lenguajes con los que lidiamos en esta tesis. En este sentido introdujimos axiomatizaciones, la noción de bisimulación para investigar el poder expresivo de una lógica, y los problemas de razonamiento asociados y su complejidad.

En el Capítulo 2, nos enfocamos en el aspecto epistémico de las lógicas modales. Es decir, en la representación del conocimiento de un conjunto de agentes. Para ello, inicialmente empezamos definiendo la lógica epistémica clásica de "saber qué", orientada a describir el conocimiento que tienen los agentes sobre ciertos hechos proposicionales [93, 108]. Para ello, los modelos de esta lógica utilizan relaciones de indistinguibilidad entre mundos para cada agente, moldeando así la percepción que tienen los mismos. Posteriormente, se hace un relevamiento sobre algunas propuestas sobre las cuales el objeto de conocimiento no son hechos proposicionales, sino eventos, justificaciones, valores o habilidades. Particularmente, esta última explora el "saber cómo" de los agentes, distinta del "saber qué", y que ha sido nuestro objeto de estudio.

En el Capítulo 3 presentamos formalmente una lógica de saber cómo introducida originalmente en [116–118], y que cuenta con un amplio consenso en la comunidad de lógica epistémica como una manera apropiada de describir el saber cómo. La lógica L_{Kh} cuenta con una modalidad binaria global $\mathsf{Kh}(\psi, \varphi)$, la cual se interpreta sobre planes lineales en sistemas de transiciones etiquetadas (LTSs), de manera similar a la lógica proposicional dinámica (PDL). En la misma, se consideran estados y acciones que al ser ejecutadas, permiten ir de un estado a otro. En la semántica, se considera que "un agente sabe cómo ir de ψ a φ" ($\mathsf{Kh}(\psi, \varphi)$) si y sólo si existe un plan σ, una serie de acciones, fuertemente ejecutable en todos los estados del LTS en donde se cumpla ψ, tal que desde estos alcance únicamente estados donde se cumplan φ. Esta restricción de ejecutabilidad fuerte significa que toda ejecución parcial de σ desde cualquier estado ψ siempre se completa. Considerando que la modalidad universal (A) es definible en términos de Kh, se presentó una axiomatización correcta y fuertemente completa, $\mathcal{L}_{\mathsf{Kh}}^{\mathrm{LTS}}$, de esta lógica. De manera análoga al Capítulo 1, definimos las nociones de bisimulaciones y equivalencia de fórmulas, utilizando conjuntos proposicionalmente definibles en vez de estados particulares, dado que Kh es una modalidad global. Además, se tienen resultados de complejidad computacional. Utilizando autómatas determinísticos finitos y lenguajes regulares, se obtiene que model checking es PSpace-completo. Mientras tanto, mediante un análisis proposicional de fórmulas, satisfacibilidad está $\mathsf{NP}^{\mathsf{NP}}$. La demostración de este resultado no fue incluida en la tesis, pero se encuentra en [6]. Finalmente, concluimos el capítulo revisando otras propuestas para modelar el saber cómo.

Aún así, esta lógica presenta ciertas dificultades. En particular, supone condiciones demasiado "ideales" para el agente. Por ejemplo, el grado de omnisciencia en los mismos. Estas condiciones pueden visualizarse en los axiomas EMP y COMP, en los cuales se asume que el agente puede transformar propiedades universales en conocimiento, y que tiene la suficiente "introspección" para poder componer planes de manera adecuada. Más aún, existen situaciones por las cuales un agente puede no saber cómo lograr un cierto objetivo, que difícilmente pueden ser representadas en la semántica de esta lógica. Por ejemplo, el agente puede no distinguir un plan "bueno" de otro que no lo es, o sólo tener un conjunto limitado de cursos de acción a considerar. Esto se debe a que en este enfoque, los planes disponibles en el modelo LTS y los que percibe el agente son los mismos. Por estos motivos, inspirados en la lógica epistémica clásica, en el Capítulo 4 definimos una lógica, $\mathsf{L}_{\mathsf{Kh}_i}$, en la cual se agrega en los LTS una relación de indistinguibilidad entre planes para cada agente, denominando a las estructuras resultantes LTS con incertidumbre, introducida inicialmente en [8, 10]. Esto permite definir un conjunto de agentes que compartan un mismo LTS, pero que a la vez difieran en la percepción que tienen del mismo, dándose así una separación explícita entre las habilidades disponibles (información óntica), de las acciones percibidas por los agentes (información epistémica). De esta manera consideramos que es posible obtener una representanción más realista de los agentes donde, por ejemplo, los axiomas EMP y COMP no sean necesariamente válidos. En la nueva semántica, "un agente i sabe cómo lograr φ a partir de ψ" ($\mathsf{Kh}_i(\psi, \varphi)$) si y sólo si existe un conjunto de planes indistinguibles entre sí para i, tal que todos los planes del conjunto son fuertemente ejecutables en los estados del LTS donde se cumple ψ, y que al ejecutar cualquiera de ellos el agente alcance únicamente estados donde se cumple φ.

El incumplimiento de los axiomas EMP y COMP no dificulta la definición de un sistema axiomático para esta lógica. Tomando de [8], en el Capítulo 5 se presenta una axiomatización similar a $\mathcal{L}^{\mathrm{LTS}}_{\mathsf{Kh}}$, $\mathcal{L}^{\mathrm{LTS}^U}_{\mathsf{Kh}_i}$, en la cual se reemplazan estos axiomas por un solo axioma más débil denominado KhA. A pesar de que esto todavía permite cierto nivel de omnisciencia por parte de los agentes, se lo considera más realista que el enfoque anterior.

En el Capítulo 6 introducimos la noción de bisimulación para esta lógica, desarrollada en [10] e inspirada en la noción correspondiente para L_{Kh}. Con estas definiciones demostramos los resultados de correspondencia esperados, es decir los teoremas de invarianza y de Hennessy-Milner sobre la clase de modelos LTS^Us cuya componente LTS es finita ($\mathbf{M_{FD}}$). Además, se demuestra que la lógica $\mathsf{L}_{\mathsf{Kh}_i}$ es más débil que L_{Kh}, mostrando que KhA es derivable en $\mathcal{L}^{\mathrm{LTS}}_{\mathsf{Kh}}$ utilizando EMP y COMP. Sin embargo, nuestra propuesta generaliza a L_{Kh} en otro sentido. Es decir, demostramos que existe una clase de modelos LTS^Us tal que cada LTS se corresponde unívocamente con un LTS^U de dicha clase y viceversa. En este trabajo logramos dar con dos de estas clases. En la primera de estas, demostrada en [8], se tiene que el agente considera todos los planes posibles y es capaz de distinguirlos entre sí ($\mathbf{M_{NU}}$). Mientras tanto, en la segunda capturamos los modelos LTS^Us que tienen una acción que simule el plan vacío ϵ (activos) y que para cada par de conjunto de planes, exista un tercero que actúe como la composición de estos (composicionales) ($\mathbf{M_{AC}}$). En el caso de esta última clase, se toma una definición ligeramente modificada de la vista en [10]. Esta construcción se debe a que necesitamos caracterizar los axiomas EMP y COMP de la propuesta original.

El Capítulo 7 está dedicado a demostrar que el problema de satisfacibilidad es decidible, y a caracterizar su complejidad, resultados demostrados en [10]. Primero consideramos un modelo arbitrario que satisfaga la fórmula en cuestión, y definimos una filtración que nos permita obtener un modelo finito. Para ello definimos una relación de equivalencia para estados que satisfagan las mismas fórmulas y otra para los planes que coinciden en ser testigos de las mismas modalidades Kh_i. Con esto, obtenemos un modelo finito, aunque de tamaño potencialmente exponencial con respecto a la fórmula misma. Notar que la definición de filtraciones es ligeramente diferente de la vista en [10]. Esto se hace con el objetivo de tener filtraciones más manejables y convenientes en la práctica, dando condiciones más simples y fáciles de verificar. Por otro lado, consideramos alguno de los modelos canónicos que satisface la fórmula y definimos una función de selección que nos permite extraer estados y planes testigos para cada Kh_i que aparece en la fórmula a tratar. Dicha selección resulta en un modelo de tamaño polinomial con respecto a la fórmula original. Además, dado que el problema de model checking está en P y que la lógica proposicional es un fragmento de $\mathsf{L}_{\mathsf{Kh}_i}$, tenemos que satisfacibilidad para $\mathsf{L}_{\mathsf{Kh}_i}$ es NP-completo.

Como se mencionó anteriormente, la introducción de la relación de indistinguibilidad entre planes para cada agente permite considerar dos tipos de información: la información óntica, los planes disponibles en la parte LTS del modelo, y la información epistémica, los planes de los cuales es consciente cada agente i en su correspondiente conjunto $\mathrm{U}(i)$. Además de tener representaciones más reales de los agentes, estos nos permite realizar un tratamiento de la lógica desde un punto de vista epistémico. En este sentido, en el Capítulo 8, basado en los resultados de [9], introducimos operadores dinámicos que modifican los diferentes tipos de información de un agente, de manera similar a las lógicas dinámicas epistémicas clásicas. Para la información óntica, se introducen dos operadores. Uno basado en anuncios públicos ($[!\chi]$ en $\mathsf{PAL}_{\mathsf{Kh}_i}$) y otro basado en actualización de relaciones ($[E]$ en $\mathsf{AUL}_{\mathsf{Kh}_i}$), que originalmente se correspondían con cambios epistémicos en la lógica epistémica clásica modificando la relación de indistinguibilidad entre mundos. Para estas modalidades que extienden $\mathsf{L}_{\mathsf{Kh}_i}$, demostramos que cada una por separado es más expresiva con respecto a la clase de todos los modelos LTS^Us. Para ellos es crucial la definición de bisimulación introducida en capítulos anteriores. Este resultado se debe a que la modalidad Kh_i no es capaz de hablar explícitamente de los planes testigo que tiene. Sin embargo, restringiéndonos a una clase específica de modelos donde los planes considerados son acciones básicas ($\mathbf{M_{BA}}$), conseguimos axiomas de reducción y, por lo tanto, una axiomatización correcta y fuertemente completa. Esta restricción tiene sentido si consideramos abstracciones menos detalladas de ciertos escenarios. Además, la lógica es decidible en esta clase utilizando el resultado del Capítulo 7 sobre la clase $\mathbf{M}^{\mathbf{F}}_{\mathbf{NU}}$. Por otro lado, para la información epistémica, introducimos tres modalidades que extienden $\mathsf{L}_{\mathsf{Kh}_i}$ y refinan la relación de indistinguibilidad de los agentes, generando particiones de un conjunto de planes determinado. La primera refinando planes específicos ($\langle \sigma \not\sim \sigma' \rangle$ en $\mathsf{L}_{\mathsf{Ref}}$), la segunda refinando planes arbitrarios ($\langle \not\sim \rangle$ en $\mathsf{L}_{\mathsf{ARef}}$) y la tercera definiendo una noción de "aprender cómo" (es decir, que el agente aprenda habilidades) utilizando el refinamiento arbitrario ($\langle \psi, \varphi \rangle_i$ en L_{Lh}). Si bien las modalidades de estas lógicas son normales, no cumplen con la propiedad de sustitución uniforme. Esto dificulta la definición de axiomas de reducción y conseguir axiomatizaciones teniendo $\mathsf{L}_{\mathsf{Kh}_i}$ como lenguaje base.

Considerando estas situaciones, un camino posible es enriquecer $\mathsf{L}_{\mathsf{Kh}_i}$ con modalidades capaces de expresar propiedades de los planes de manera explícita. Para ello, en el Capítulo 9 decidimos enriquecer el lenguaje base con modalidades $\Box$ de la lógica multimodal básica, en este caso indexadas por acciones ($\mathsf{L}_{\mathsf{Kh}_i,\Box}$). Con esto, demostramos que la lógica tiene una axiomatización correcta y fuertemente completa. Además, utilizando una noción de filtraciones similar a la vista en el Capítulo 7, se demostró que el problema de satisfacibilidad es decidible. Con este nuevo lenguaje como base, se presentan tres operadores dinámicos epistémicos que hacen que el agente sea capaz de distinguir un plan específico de todos los demás, independientemente si lo consideraba o no. Las primeras dos extensiones ($[!a]$ en $\mathsf{L}_{\mathsf{Kh}_i,\Box,[!a]}$ y $[!\sigma]$ en $\mathsf{L}_{\mathsf{Kh}_i,\Box,[!\sigma]}$) son de carácter público. Es decir, se anuncia a todos los agentes que un determinado plan es distinguible de otros. Mientras tanto, la tercera definimos un anuncio privado, es decir, que revela información a un agente específico ($[!\sigma, i]$ en $\mathsf{L}_{\mathsf{Kh}_i,\Box,[!\sigma,i]}$). Para los tres operadores definimos axiomas de reducción a $\mathsf{L}_{\mathsf{Kh}_i,\Box}$ sobre la clase de todos los LTS^Us. Por lo tanto, se tiene una axiomatización correcta y fuertemente completa, y que los problemas de satisfacibilidad para cada extensión son decidibles.

Finalmente, en el Capítulo 10, de la misma manera que la lógica modal básica puede tener diferentes interpretaciones, en nuestro caso reinterpretamos la modalidad de saber cómo Kh_i en un entorno deóntico. En este contexto introducimos tres modalidades que nos permiten razonar acerca de ***(1)*** las habilidades de los agentes ($\mathsf{S}(\psi, \varphi)$), ***(2)*** las normas con las que deben cumplir ($\mathsf{N}(\psi, \varphi)$) y ***(3)*** sobre cómo los agentes *cumplen conscientemente* o no con estas ($\mathsf{Kc}_i(\psi, \varphi)$). Para esto, extendimos nuestro modelos con un conjunto de normas, definiendo así los llamados LTS^Us normativos (NLTS^Us). En la parte semántica, S se comporta exactamenmte como la modalidad Kh en L_{Kh}, N restringe el rango de planes de Act^* a los del conjunto N y Kc_i se comporta como Kh_i con la salvedad de que el conjunto de planes necesariamente debe estar incluido en el conjunto de normas. Es decir, un agente cumple conscientemente si el conjunto de normas cumple con las condiciones originales de ejecutabilidad fuerte y alcanzabilidad, y además dicho conjunto está permitido por las normas. Por una parte, estudiamos los fragmentos $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N}}$ y $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N},\mathsf{S}}$ por separado. En ambos se consiguieron axiomatizaciones correctas y fuertemente completas. Más aún, para $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N}}$, utilizando la función de selección del Capítulo 7 para $\mathsf{L}_{\mathsf{Kh}_i}$, el problema de model checking está en P y satisfacibilidad es NP-completo. Estos resultados fueron demostrados en [5].

Capítulo 12

Trabajo futuro

> I cannot escape from the conclusion that the great ages of progress have depended upon a small number of individuals of transcendent ability.
>
> Bertrand Russell

En este trabajo hemos discutido y demostrado propiedades relevantes alrededor de la lógica $\mathsf{L_{Kh_i}}$. Sin embargo, quedan muchos resultados a ser investigados en el futuro. Discutiremos algunos de ellos en este capítulo.

Una de las preguntas que surgen es si la noción de ejecutabilidad fuerte es necesaria para modelar este tipo de conocimiento. Por ejemplo, es posible considerar los planes que son débilmente ejecutables. Es decir, aquellos que pueden fallar en algunas ejecuciones parciales pero que al menos existe una que logra terminar, o considerar todos los puntos en donde terminan las ejecuciones parciales, sean fallidas o no. Más aún, dada la reinterpretación deóntica del Capítulo 10, el agregar la restricción $\pi \subseteq \mathrm{N}$ abre la posibilidad de definir nuevos operadores que dependan de nuevas componentes en los modelos y las restricciones que consideremos convenientes imponer. Por ejemplo, las normas que ciertos agentes deben cumplir pero que otros no necesariamente, como una colección de normas indexada por cada agente ($\mathrm{N}(i)$), o considerar que los planes que definen el conocimiento sean no disjuntos de las normas, pero no necesariamente un subconjunto de ellas. Resulta interesante investigar las lógicas resultantes de estas variantes.

Otro de los aspectos a investigar es la clase exacta de complejidad a la que pertenece el problema de satisfacibilidad para $\mathsf{L_{Kh}}$. Si bien sabemos que la cota superior es $\mathsf{NP^{NP}}$ (resultado no presentado en detalle en esta tesis), para demostrar que es $\mathsf{NP^{NP}}$-completo, es necesario encontrar un problema equivalente en $\mathsf{NP^{NP}}$ (por ejemplo, Truth of Quantified Boolean Formula [12]) que pueda ser reducido polinomialmente a este. Aún así, dado que la satisfacibilidad de una fórmula en $\mathsf{L_{Kh}}$ se reduce a analizar la satisfacibilidad de un conjunto de fórmulas proposicionales, este resultado se puede utilizar en la implementación de un SAT solver específico.

Por el lado de los operadores dinámicos, queda como trabajo futuro estudiar si se pueden definir nuevas modalidades para las cuales sea posible definir axiomas de reducción, ya sea para $\mathsf{L_{Kh_i}}$ o su versión extendida con operadores $[a]$ de $\mathsf{L_{Kh_i,\Box}}$. Por ejemplo, si bien los operadores $\langle\psi,\varphi\rangle_i$ y $\mathsf{L}_i(\psi,\varphi)$ están definidos con una única modalidad $\langle\not\sim\rangle$, se podría considerar como alternativa redefinirlas como

$$\langle\psi,\varphi\rangle_i^*\chi := \bigvee_{n\geq 1}\langle\not\sim\rangle^n(\mathsf{Kh}_i(\psi,\varphi)\wedge\chi),$$

y $\mathsf{L}_i^*(\psi,\varphi) := \langle\psi,\varphi\rangle_i^*\top$, siendo $\langle\not\sim\rangle^n$ n veces $\langle\not\sim\rangle$. Ambos operadores describirían la capacidad que el agente i tiene de aprender "eventualmente" una determinada habilidad, dados los suficientes pasos de "aprendizaje", en vez de considerar sólo uno de estos. Otra alternativa es definir un operador que emule la acción de "olvidar cómo", ya sea mediante la eliminación de un plan o un conjunto de planes de un agente, o la confusión/fusión con otro. Además, queda como tarea a futuro investigar si existen otras clases de modelos interesantes de las cuales se pueden obtener axiomas de reducción o, axiomatizaciones sin la necesidad de contar con sustitución uniforme. También resta saber si $\mathsf{L}_{\mathsf{Ref}}$ y $\mathsf{L}_{\mathsf{ARef}}$ son comparables en poder expresivo o no. Si bien, como consecuencia de la Definición 8.15, $\models \langle \sigma_1 \not\sim \sigma_2 \rangle \varphi \rightarrow \langle \not\sim \rangle \varphi$, caracterizar la relación de expresividad entre las dos lógicas requerirá más desarrollo. Particularmente porque $\mathsf{L}_{\mathsf{Ref}}$ es capaz de hablar sobre planes específicos mientras que no es el caso para $\mathsf{L}_{\mathsf{ARef}}$. Por otro lado, si bien se ha demostrado en algunos casos la decidibilidad de los problemas de satisfacibilidad de algunas extensiones como $\mathsf{L}_{\mathsf{Kh}_i,\Box}$, resta analizar cuál sería la complejidad computacional exacta de estas. Por ejemplo, para $\mathsf{PAL}_{\mathsf{Kh}_i}$ se puede estimar que es a lo sumo $\mathsf{ExpTime}$, aunque requeriría más estudio saber si pertenece a una clase menor.

Volviendo nuevamente a la interpretación deóntica, queda abierta la caracterización exacta de la la complejidad de la lógica $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N},\mathsf{S}}$, que incluye a los tres operadores. Dado que $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N}}$ es NP-completo y L_{S} es a lo sumo $\mathsf{NP}^{\mathsf{NP}}$ (S y Kh tienen la misma semántica), podemos conjeturar que $\mathsf{L}_{\mathsf{Kc}_i,\mathsf{N},\mathsf{S}}$ también es a lo sumo $\mathsf{NP}^{\mathsf{NP}}$.

Bibliografía

[1] Thomas Ågotnes, Philippe Balbiani, Hans van Ditmarsch, and Pablo Seban. Group announcement logic. *Journal of Applied Logic*, 8(1):62–81, 2009. [cited in page(s): 98]

[2] Thomas Ågotnes and Hans van Ditmarsch. Coalitions and announcements. In *7th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2008), Volume 2*, pages 673–680. IFAAMAS, 2008. [cited in page(s): 98]

[3] Rajeev Alur, Thomas A. Henzinger, and Orna Kupferman. Alternating-time temporal logic. In *38th Annual Symposium on Foundations of Computer Science, FOCS '97, Miami Beach, Florida, USA, October 19-22, 1997*, pages 100–109. IEEE Computer Society, 1997. [cited in page(s): 19]

[4] Leenart Åqvist. Deontic logic. In Dov M. Gabbay and Franz Guenthner, editors, *Handbook of Philosophical Logic: Volume 8*, pages 147–264. Springer Netherlands, Dordrecht, 2002. [cited in page(s): 117]

[5] Carlos Areces, Valentin Cassano, Pablo F. Castro, Raul Fervari, and Andrés R. Saravia. A deontic logic of knowingly complying. In Noa Agmon, Bo An, Alessandro Ricci, and William Yeoh, editors, *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2023, London, United Kingdom, 29 May 2023 - 2 June 2023*, pages 364–372. ACM, 2023. [cited in page(s): 32, 33, 118, 122, 124, 126, 142]

[6] Carlos Areces, Valentin Cassano, Pablo F. Castro, Raul Fervari, and Andrés R. Saravia. How easy it is to know how: An upper bound for the satisfiability problem. In *18th Edition of the European Conference on Logics in Artificial Intelligence (JELIA 2023)*, 2023. [cited in page(s): 32, 46, 140]

[7] Carlos Areces, Raul Fervari, and Guillaume Hoffmann. Relation-changing modal operators. *Log. J. IGPL*, 23(4):601–627, 2015. [cited in page(s): 29, 85, 98]

[8] Carlos Areces, Raul Fervari, Andrés R. Saravia, and Fernando R. Velázquez-Quesada. Uncertainty-based semantics for multi-agent knowing how logics. In *18th Conference on Theoretical Aspects of Rationality and Knowledge (TARK 2021)*, volume 335 of *EPTCS*, pages 23–37. Open Publishing Association, 2021. [cited in page(s): 32, 48, 53, 55, 131, 140]

[9] Carlos Areces, Raul Fervari, Andrés R. Saravia, and Fernando R. Velázquez-Quesada. First steps in updating knowing how. In Carlos Areces and Diana Costa, editors, *Dynamic Logic. New Trends and Applications*, pages 1–16. Springer International Publishing, 2023. [cited in page(s): 32, 59, 86, 103, 141]

[10] Carlos Areces, Raul Fervari, Andrés R. Saravia, and Fernando R. Velázquez-Quesada. Uncertainty-based knowing how logic. *Journal of Logic and Computation*, page exad056, 10 2023. [cited in page(s): 32, 48, 53, 55, 61, 68, 75, 76, 140]

[11] Carlos Areces, Hans van Ditmarsch, Raul Fervari, Bastien Maubert, and François Schwarzentruber. Copy and remove as dynamic operators. *J. Appl. Non Class. Logics*, 31(3-4):181–220, 2021. [cited in page(s): 24]

[12] Sanjeev Arora and Boaz Barak. *Computational Complexity: A Modern Approach.* Cambridge University Press, 2006. [cited in page(s): 26, 45, 46, 81, 131, 143]

[13] Sergei Artemov. The logic of justification. *The Review of Symbolic Logic*, 1(04):477–513, 2008. [cited in page(s): 31]

[14] Philippe Balbiani, Alexandru Baltag, Hans van Ditmarsch, Andreas Herzig, Tomohiro Hoshi, and Tiago de Lima. 'knowable' as 'known after an announcement'. *The Review of Symbolic Logic*, 1(3):305–334, 2008. [cited in page(s): 98]

[15] Philippe Balbiani, Andreas Herzig, and Nicolas Troquard. Alternative axiomatics and complexity of deliberative stit theories. *Journal of Philosophical Logic*, 37(4):387–406, 2008. [cited in page(s): 118]

[16] Alexandru Baltag. To know is to know the value of a variable. In *Advances in Modal Logic 11*, pages 135–155, 2016. [cited in page(s): 31]

[17] Alexandru Baltag, Lawrence S. Moss, and Slawomir Solecki. The logic of public announcements, common knowledge, and private suspicions. In *Proceedings of TARK '98*, pages 43–56. Morgan Kaufmann Publishers Inc., 1998. [cited in page(s): 86]

[18] Alexandru Baltag, Lawrence S. Moss, and Sławomir Solecki. *The Logic of Public Announcements, Common Knowledge, and Private Suspicions*, pages 773–812. Springer Graduate Texts in Philosophy 1. Springer International Publishing, Cham, 1 edition, 2016. [cited in page(s): 86]

[19] Alexandru Baltag and Bryan Renne. Dynamic epistemic logic. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2016 edition, 2016. [cited in page(s): 85, 114]

[20] Oskar Becker. *Untersuchungen Über den Modalkalkül*. A. Hain, 1952. [cited in page(s): 117]

[21] Francesco Belardinelli. Reasoning about knowledge and strategies: Epistemic strategy logic. In Fabio Mogavero, Aniello Murano, and Moshe Y. Vardi, editors, *Proceedings 2nd International Workshop on Strategic Reasoning, SR 2014, Grenoble, France, April 5-6, 2014*, volume 146 of *EPTCS*, pages 27–33, 2014. [cited in page(s): 32]

[22] Nuel Belnap and Michael Perloff. Seeing to it that: A canonical form for agentives. *Theoria*, 54 (3):175–199, 1988. [cited in page(s): 117]

[23] Martin Mose Bentzen. *Stit, Iit, and Deontic logic for Action Types*. PhD thesis, Section for Philsophy and Science Studies, Roskilde University, 2010. [cited in page(s): 33, 118]

[24] Patrick Blackburn, Maarten de Rijke, and Yde Venema. *Modal Logic*. Cambridge University Press, November 2002. [cited in page(s): 15, 16, 20, 21, 22, 28, 29, 55, 56, 75, 76, 81, 103, 108, 109, 139]

[25] Patrick Blackburn and Johan F.A.K. van Benthem. Modal logic: A semantic perspective. In *Handbook of Modal Logic*, pages 1–84. Elsevier, 2006. [cited in page(s): 26, 46]

[26] Patrick Blackburn, Johan F.A.K. van Benthem, and Frank Wolter Blackburn. *Handbook of Modal Logic*, volume 3. Elsevier Science Inc., New York, NY, USA, 2006. [cited in page(s): 15, 16, 17, 20, 21, 22, 24, 28, 29, 30, 31, 45, 80, 103, 108, 109, 130, 139]

[27] Jan Broersen. A logical analysis of the interaction between 'obligation-to-do' and 'knowingly doing'. In *Proceedings of DEON 08*, pages 140–154, 2008. [cited in page(s): 117]

[28] Jan Broersen. Deontic epistemic stit logic distinguishing modes of mens rea. *Journal of Applied Logic*, 9(2):137–152, 2011. [cited in page(s): 117]

[29] Jan Broersen. Making a start with the *stit* logic analysis of intentional action. *Journal of Philosophical Logic*, 40(4):499–530, 2011. [cited in page(s): 117]

[30] Brian F. Chellas. *Modal Logic: An Introduction*. Cambridge University Press, 1980. [cited in page(s): 15]

[31] Amit Chopra, Leendert van der Torre, Harko Verhagen, and Serena Villata, editors. *Handbook of Normative Multiagent Systems*. College Publications, 2018. [cited in page(s): 117]

[32] Kim Guldstrand Larsen Christel Baier, Joost-Pieter Katoen. *Principles of Model Checking*. MiT Press, 2008. [cited in page(s): 17]

[33] Edmund M. Clarke and E. Allen Emerson. Design and synthesis of synchronization skeletons using branching time temporal logic. In Dexter Kozen, editor, *Logics of Programs*, pages 52–71, Berlin, Heidelberg, 1982. Springer Berlin Heidelberg. [cited in page(s): 17]

[34] Mika Cohen and Mads Dam. Logical omniscience in the semantics of BAN logic. In *Proceedings of FCS '05*, 2005. [cited in page(s): 53]

[35] Stephen A. Cook. The complexity of theorem-proving procedures. In Michael A. Harrison, Ranan B. Banerji, and Jeffrey D. Ullman, editors, *Proceedings of the 3rd Annual ACM Symposium on Theory of Computing, May 3-5, 1971, Shaker Heights, Ohio, USA*, pages 151–158. ACM, 1971. [cited in page(s): 26, 45, 81, 131]

[36] Stéphane Demri and Raul Fervari. Model-checking for ability-based logics with constrained plans. In *37th AAAI Conference on Artificial Intelligence (AAAI 2023)*, pages 6305–6312. AAAI Press, 2023. [cited in page(s): 32, 46]

[37] Herbert B. Enderton. *A Mathematical Introduction to Logic, Second Edition*. Academic Press, 2 edition, 2001. [cited in page(s): 16]

[38] Ronald Fagin and Joseph Y. Halpern. Belief, awareness, and limited reasoning. *Artificial Intelligence*, 34(1):39–76, 1988. [cited in page(s): 30, 32]

[39] Ronald Fagin, Joseph Y. Halpern, Yoram Moses, and Moshe Y. Vardi. *Reasoning about knowledge*. The MIT Press, Cambridge, Mass., 1995. [cited in page(s): 27, 30, 85]

[40] Jie Fan, Yanjing Wang, and Hans van Ditmarsch. Contingency and knowing whether. *The Review of Symbolic Logic*, 8(1):75–107, 2015. [cited in page(s): 31]

[41] Jeremy Fantl. Knowledge how. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2021 edition, 2021. [cited in page(s): 19, 31]

[42] Raul Fervari, Andreas Herzig, Yanjun Li, and Yanjing Wang. Strategically knowing how. In *26th International Joint Conference on Artificial Intelligence (IJCAI 2017)*, pages 1031–1038. International Joint Conferences on Artificial Intelligence, 2017. [cited in page(s): 31, 32, 46]

[43] Raul Fervari, Fernando R. Velázquez-Quesada, and Yanjing Wang. Bisimulations for *Knowing How* logics. In *5th International Workshop on Strategic Reasoning, SR 2017*, 2017. [cited in page(s): 32, 42, 43, 44, 45, 61]

[44] Raul Fervari, Fernando R. Velázquez-Quesada, and Yanjing Wang. Bisimulations for *Knowing How* logics. *The Review of Symbolic Logic*, 15(2):450–486, 2022. [cited in page(s): 32, 42, 43, 44, 45, 46, 61, 64]

[45] Michael J. Fischer and Richard E. Ladner. Propositional dynamic logic of regular programs. *Journal of Computer and System Sciences*, 18(2):194–211, 1979. [cited in page(s): 17, 31]

[46] Dov Gabbay, John Horty, Xavier Parent, Ron van der Meyden, and Leendert van der Torre, editors. *Handbook of Deontic Logic and Normative Systems*, volume 1. College Publications, 2013. [cited in page(s): 117]

[47] Dov Gabbay, John Horty, Xavier Parent, Ron van der Meyden, and Leendert van der Torre, editors. *Handbook of Deontic Logic and Normative Systems*, volume 2. College Publications, 2021. [cited in page(s): 117]

[48] James Garson. Modal logic. In Edward N. Zalta and Uri Nodelman, editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2023 edition, 2023. [cited in page(s): 21, 22]

[49] Malik Ghallab, Dana S. Nau, and Paolo Traverso. *Automated Planning and Acting*. Cambridge University Press, 2016. [cited in page(s): 19, 31, 90]

[50] Robert Goldblatt. Mathematical modal logic: A view of its evolution. *J. Appl. Log.*, 1(5-6):309–392, 2003. [cited in page(s): 15]

[51] Valentin Goranko and Solomon Passy. Using the universal modality: Gains and questions. *Journal of Logic and Computation*, 2(1):5–30, 1992. [cited in page(s): 41, 119]

[52] Tao Gu and Yanjing Wang. "Knowing value" logic as a normal modal logic. In Lev D. Beklemishev, Stéphane Demri, and András Maté, editors, *Advances in Modal Logic 11, proceedings of the 11th conference on .^Advances in Modal Logic,"held in Budapest, Hungary, August 30 - September 2, 2016*, pages 362–381. College Publications, 2016. [cited in page(s): 31]

[53] Joseph Y. Halpern and Yoram Moses. Knowledge and common knowledge in a distributed environment. *Journal of the ACM*, 37(3):549–587, 1990. [cited in page(s): 85]

[54] Joseph Y. Halpern and Riccardo Pucella. Dealing with logical omniscience: Expressiveness and pragmatics. *Artificial Intelligence*, 175(1):220–235, 2011. [cited in page(s): 32, 53]

[55] David Harel, Dexter Kozen, and Jerzy Tiuryn. *Dynamic Logic*. The MIT Press, 2000. [cited in page(s): 17, 31]

[56] Sergiu Hart, Aviad Heifetz, and Dov Samet. Knowing whether, knowing that, and the cardinality of state spaces. *Journal of Economic Theory*, 70(1):249–256, 1996. [cited in page(s): 31]

[57] Katherine Hawley. Success and knowledge-how. *American Philosophical Quarterly*, 40(1):19–31, 2003. [cited in page(s): 32]

[58] Vincent F. Hendricks. Introduction: 8 bridges between mainstream and formal epistemology. *Philosophical Studies*, 128(1):1–5, March 2006. [cited in page(s): 27]

[59] Andreas Herzig. Logics of knowledge and action: critical analysis and challenges. *Autonomous Agents and Multi Agent Systems*, 29(5):719–753, 2015. [cited in page(s): 20, 31]

[60] Andreas Herzig and François Schwarzentruber. Properties of logics of individual and group agency. In Carlos Areces and Robert Goldblatt, editors, *Advances in Modal Logic 7, Nancy, France, 9-12 September 2008*, pages 133–149. College Publications, 2008. [cited in page(s): 118]

[61] Andreas Herzig and Nicolas Troquard. Knowing how to play: uniform choices in logics of agency. In Hideyuki Nakashima, Michael P. Wellman, Gerhard Weiss, and Peter Stone, editors, *5th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2006), Hakodate, Japan, May 8-12, 2006*, pages 209–216. ACM, 2006. [cited in page(s): 31]

[62] Jaakko Hintikka. *Knowledge and Belief*. Cornell University Press, Ithaca, NY, 1962. [cited in page(s): 19, 27, 30, 31, 48]

[63] Wesley H. Holliday, Tomohiro Hoshi, and Thomas F. Icard III. Schematic validity in dynamic epistemic logic: Decidability. In *Proceedings of LORI-3*, 2011. [cited in page(s): 22]

[64] John F. Horty. *Agency and Deontic Logic*. Oxford University Press, 2001. [cited in page(s): 117, 118]

[65] Wojciech Jamroga. *Logical methods for specification and verification of multi-agent systems*. Institute of Computer Science, Polish Academy of Sciences, 2015. [cited in page(s): 18]

[66] Wojciech Jamroga and Thomas Ågotnes. Constructive knowledge: what agents can achieve under imperfect information. *Journal of Applied Non Classical Logics*, 17(4):423–475, 2007. [cited in page(s): 20, 31]

[67] Wojciech Jamroga and Wiebe van der Hoek. Agents that know how to play. *Fundamenta Informaticae*, 63(2-3):185–219, 2004. [cited in page(s): 32]

[68] Jerzy Kalinowski. Theorie des propositions normatives. *Studia Logica*, 1(1):147–182, 1953. [cited in page(s): 117]

[69] Barteld Kooi and Bryan Renne. Arrow update logic. *The Review of Symbolic Logic*, 4(4):536–559, 2011. [cited in page(s): 32, 85, 86, 92]

[70] E. J. Lemmon and Dana Scott. Intensional logic, preliminary draft of initial chapters by ej lemmon, july 1966, nowadays available as an introduction to modal logic (american philosophical quarterly monograph no. 11) edited by k. segerberg. *American Philosophical Quarterly Monograph*, 11, 1977. [cited in page(s): 75]

[71] Yves Lespérance, Hector J. Levesque, Fangzhen Lin, and Richard B. Scherl. Ability and knowing how in the situation calculus. *Studia Logica*, 66(1):165–186, 2000. [cited in page(s): 31]

[72] Yanjun Li. *Knowing what to do: a logical approach to planning and knowing how*. PhD thesis, University of Groningen, 2017. [cited in page(s): 19, 32, 46, 134]

[73] Yanjun Li. Stopping means achieving: A weaker logic of knowing how. *Studies in Logic*, 9(4):34–54, 2017. [cited in page(s): 31, 32, 38, 46]

[74] Yanjun Li. Tableau-based decision procedure for logic of knowing-how via simple plans. In Pietro Baroni, Christoph Benzmüller, and Yì N. Wáng, editors, *Logic and Argumentation - 4th International Conference, CLAR 2021, Hangzhou, China, October 20-22, 2021, Proceedings*, volume 13040 of *Lecture Notes in Computer Science*, pages 266–283. Springer, 2021. [cited in page(s): 46]

[75] Yanjun Li. Tableaux for the logic of strategically knowing how. In *19th Conference on Theoretical Aspects of Rationality and Knowledge (TARK 2023)*, volume 379 of *EPTCS*, pages 379–391. Open Publishing Association, 2023. [cited in page(s): 46]

[76] Yanjun Li and Yanjing Wang. Achieving while maintaining: A logic of knowing how with intermediate constraints. In *Logic and Its Applications - 7th Indian Conference, ICLA 2017*, pages 154–167, 2017. [cited in page(s): 31, 32, 38, 46]

[77] Yanjun Li and Yanjing Wang. Planning-based knowing how: A unified approach. *Artificial Intelligence*, 296, 2021. [cited in page(s): 46]

[78] Emiliano Lorini. Temporal stit logic and its application to normative reasoning. *Journal of Applied Non-Classical Logics*, 23(4):372–399, 2013. [cited in page(s): 117]

[79] Minghui Ma and Katsuhiko Sano. How to update neighbourhood models. *Journal of Logic and Computation*, 28(8):1781–1804, 2018. [cited in page(s): 89]

[80] John McCarthy and Patrick J. Hayes. Some philosophical problems from the standpoint of artificial intelligence. In *Machine Intelligence*, pages 463–502. Edinburgh University Press, 1969. [cited in page(s): 31]

[81] John-Jules Ch. Meyer and Wiebe van der Hoek. *Epistemic Logic for AI and Computer Science*. Cambridge University Press, New York, N.Y., U.S.A., 1995. [cited in page(s): 27]

[82] Richard Montague. Universal grammar. *Theoria*, 36(3):373–398, 1970. [cited in page(s): 89]

[83] Robert Moore. A formal theory of knowledge and action. In *Formal Theories of the Commonsense World*. Ablex Publishing Corporation, 1985. [cited in page(s): 31]

[84] Antonio Moreno. Avoiding logical omniscience and perfect reasoning: a survey. *AI Communications*, 11(2):101–122, January 1998. [cited in page(s): 53]

[85] Pavel Naumov and Jia Tao. Together we know how to achieve: An epistemic logic of know-how (extended abstract). In Jérôme Lang, editor, *Proceedings Sixteenth Conference on Theoretical Aspects of Rationality and Knowledge, TARK 2017, Liverpool, UK, 24-26 July 2017*, volume 251 of *EPTCS*, pages 441–453, 2017. [cited in page(s): 46]

[86] Pavel Naumov and Jia Tao. Second-order know-how strategies. In *17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018)*, pages 390–398. ACM, 2018. [cited in page(s): 46]

[87] Pavel Naumov and Jia Tao. Together we know how to achieve: An epistemic logic of know-how. *Artificial Intelligence*, 262:279–300, 2018. [cited in page(s): 46]

[88] Marc Pauly. *Logic for social software*. PhD thesis, Universiteit van Amsterdam, 2001. [cited in page(s): 18]

[89] Carlotta Pavese. Knowledge how. In Edward N. Zalta and Uri Nodelman, editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2022 edition, 2022. [cited in page(s): 31]

[90] Andrés Perea. *Epistemic Game Theory: Reasoning and Choice*. Cambridge University Press, Cambridge, 2012. [cited in page(s): 27]

[91] Jan Plaza. Logics of public communications. In *Proceedings of the 4th International Symposium on Methodologies for Intelligent Systems*, pages 201–216, 1989. [cited in page(s): 29, 32, 85, 86, 114]

[92] Amir Pnueli. The temporal logic of programs. In *18th Annual Symposium on Foundations of Computer Science, Providence, Rhode Island, USA, 31 October - 1 November 1977*, pages 46–57. IEEE Computer Society, 1977. [cited in page(s): 17]

[93] Rasmus Rendsvig, John Symons, and Yanjing Wang. Epistemic logic. In Edward N. Zalta and Uri Nodelman, editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2023 edition, 2023. [cited in page(s): 27, 85, 139]

[94] Gilbert Ryle. *The Concept of Mind: 60th Anniversary Edition*. Hutchinson & Co, 1949. [cited in page(s): 31]

[95] Davide Sangiorgi. On the origins of bisimulation and coinduction. *ACM Transactions on Programming Languages and Systems*, 31(4):1–41, May 2009. [cited in page(s): 23, 42]

[96] Davide Sangiorgi. *Introduction to Bisimulation and Coinduction*. Cambridge University Press, 2011. [cited in page(s): 23, 42]

[97] Dana Scott. Advice on modal logic. In Karel Lambert, editor, *Philosophical Problems in Logic*, pages 143–173. Reidel, Dordrecht, The Netherlands, 1970. [cited in page(s): 89]

[98] Krister Segerberg. Decidability of four modal logics. *Theoria*, 34(1):21–25, 1968. [cited in page(s): 75]

[99] Krister Segerberg. Decidability of s4.1. *Theoria*, 34(1):7–20, 1968. [cited in page(s): 75]

[100] David E. Smith and Daniel S. Weld. Conformant graphplan. In Jack Mostow and Chuck Rich, editors, *Proceedings of the Fifteenth National Conference on Artificial Intelligence and Tenth Innovative Applications of Artificial Intelligence Conference, AAAI 98, IAAI 98, July 26-30, 1998, Madison, Wisconsin, USA*, pages 889–896. AAAI Press / The MIT Press, 1998. [cited in page(s): 40]

[101] Robin Smith. Aristotle's logic. In Edward N. Zalta and Uri Nodelman, editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2022 edition, 2022. [cited in page(s): 15]

[102] Robert Stalnaker. The problem of logical omniscience, i. *Synthese*, 89(3):425–440, 1991. [cited in page(s): 32, 53]

[103] Jason Stanley and Timothy Williamson. Knowing how. *The Journal of Philosophy*, 98(8):411–444, 2001. [cited in page(s): 31]

[104] Nicolas Troquard and Philippe Balbiani. Propositional dynamic logic. In Edward N. Zalta and Uri Nodelman, editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2023 edition, 2023. [cited in page(s): 17, 31]

[105] Johan van Benthem. *Modal Logic for Open Minds*. Center for the Study of Language and Information - CSLI Lecture Notes 199. CSLI Publications, 2000. [cited in page(s): 15]

[106] Johan van Benthem. *Logical Dynamics of Information and Interaction*. Cambridge University Press, 2011. [cited in page(s): 22, 30]

[107] Wiebe van der Hoek, Bernd van Linder, and John-Jules Ch. Meyer. On agents that have the ability to choose. *Studia Logica*, 66(1):79–119, 2000. [cited in page(s): 31]

[108] Hans van Ditmarsch, Wiebe van der Hoek, Joseph Y. Halpern, and Barteld Kooi, editors. *Handbook of Epistemic Logic*. College Publications, 2015. [cited in page(s): 27, 85, 139]

[109] Hans van Ditmarsch, Wiebe van der Hoek, and Barteld Kooi. *Dynamic Epistemic Logic*. Springer, 2007. [cited in page(s): 29, 30, 85, 86, 88, 95]

[110] Hans van Ditmarsch, Wiebe van der Hoek, Barteld Kooi, and Louwe B. Kuijer. Arbitrary arrow update logic. *Artificial Intelligence*, 242:80–106, 2017. [cited in page(s): 86, 98]

[111] Jan van Eijck, Malvin Gattinger, and Yanjing Wang. Knowing values and public inspection. In Sujata Ghosh and Sanjiva Prasad, editors, *Logic and Its Applications - 7th Indian Conference, ICLA 2017, Kanpur, India, January 5-7, 2017, Proceedings*, volume 10119 of *Lecture Notes in Computer Science*, pages 77–90. Springer, 2017. [cited in page(s): 31]

[112] Moshe Y. Vardi. On epistemic logic and logical omniscience. In Joseph Y. Halpern, editor, *TARK*, pages 293–305. Morgan Kaufmann, 1986. [cited in page(s): 32, 53]

[113] George von Wright. Deontic logic. *Mind*, 60, 1951. [cited in page(s): 15, 16, 117]

[114] George von Wright. Deontic logic: A personal view. *Ratio Juris*, 12(1):26–38, 1999. [cited in page(s): 117]

[115] Xun Wang. A logic of knowing how with skippable plans. In *Logic, Rationality, and Interaction – 7th International Workshop, LORI 2019*, pages 413–424, 2019. [cited in page(s): 31, 32, 46]

[116] Yanjing Wang. A logic of knowing how. In *Logic, Rationality, and Interaction – 5th International Workshop, LORI 2015*, pages 392–405, 2015. [cited in page(s): 20, 31, 32, 37, 38, 39, 40, 41, 42, 53, 54, 86, 131, 134, 139]

[117] Yanjing Wang. Beyond knowing that: a new generation of epistemic logics. In Hans van Ditmarsch and Gabriel Sandu, editors, *Jaakko Hintikka on knowledge and game theoretical semantics*, pages 499–533. Springer, 2018. [cited in page(s): 20, 31, 32, 37, 38, 39, 40, 53, 131, 134, 139]

[118] Yanjing Wang. A logic of goal-directed knowing how. *Synthese*, 195(10):4419–4439, 2018. [cited in page(s): 31, 32, 37, 38, 39, 40, 53, 54, 57, 68, 73, 87, 139]

[119] Chao Xu, Yanjing Wang, and Thomas Studer. A logic of knowing why. *Synthese*, 198(2):1259–1285, 2021. [cited in page(s): 31]